Co-inoculation with novel nodule-inhabiting bacteria reduces the benefits of legume-rhizobium symbiosis

James C. Kosmopoulos[1,2,5,6,*], Rebecca T. Batstone-Doyle[2,3,4,*], Katy D. Heath[2,3]

[1]School of Integrative Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[2]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[3]Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

[4]Department of Biology, McMaster University, Hamilton, ON, Canada

[5]Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA

[6]Microbiology Doctoral Training Program, University of Wisconsin-Madison, WI, USA

[*]Address correspondence to kosmopoulos [at] wisc [dot] edu or batstonr [at] mcmaster [dot] ca

ORCIDs:

J.C.K. - 0000-0001-7070-5418

R.B.D. - 0000-0002-2530-8998

K.D.H. - 0000-0002-6368-744X

ABSTRACT: The ecologically and economically vital symbiosis between nitrogen-fixing rhizobia and leguminous plants is often thought of as a bi-partite interaction, yet studies increasingly show the prevalence of non-rhizobia endophytes (NREs) that occupy nodules alongside rhizobia. Yet, what impact these NREs have on the symbiosis remains unclear. Here, we investigated four NRE strains found to naturally co-occupy nodules with the rhizobium *Sinorhizobium meliloti* in native soils of the legume *Medicago truncatula*. Our objectives were to (1) obtain a taxonomic assignment of the NREs, (2) examine the direct and indirect effects of NREs on *M. truncatula* and *S. meliloti* fitness, and (3), determine whether NREs can re-colonize root and nodule tissues upon reinoculation. We identified two NRE strains (522 and 717A) as novel *Paenibacillus* species and the other two (702A and 733B) as novel *Pseudomonas* species. Additionally, we found that two NREs (Paenibacillus 717A and Pseudomonas 733B) reduced the fitness benefits obtained from symbiosis for both partners, while the other two (522, 702A) had little effect. Lastly, we found that NREs were able to co-infect host tissues alongside *S. meliloti*. This study demonstrates that variation of NREs present in natural populations must be considered to better understand legume-rhizobium dynamics in soil communities.

KEYWORDS: Mutualism-parasitism continuum, soil microbiome, phylogenomics, greenhouse, colonization, differential abundance

INTRODUCTION

Plant-microbe symbioses have often been studied as pairwise interactions between a single plant species and a single species of bacteria or fungi (Bakker et al. 2014; Tsiknia et al. 2020; Afkhami et al. 2020). Yet, natural microbial communities are diverse, with functional roles, assembly processes, and evolutionary dynamics that we are only beginning to understand thanks to a recent boom in plant microbiome studies (Kent and Triplett 2002; Wagner et al. 2014, 2016; Zilles et al. 2016; O'Brien et al. 2021). Thus, plants in nature never interact with a single species at a time – instead our traditional models of intimate symbiotic interactions (including plant-fungi or plant-rhizobium interactions) are occurring in a complex web of microbe-microbe (and other) interactions (Bakker et al. 2014). It has long been recognized that the impacts symbiotic partners have on one another are context-dependent, varying from beneficial (mutualistic), to commensal, to even harmful (parasitic) depending on abiotic and biotic contexts (Johnson et al. 1997; Heath and Tiffin 2007; Haney et al. 2015; Klein et al. 2022; Batstone et al. 2022b). Understanding how our best-known model symbioses are shaped by the surrounding microbial players is an important step towards uniting concepts of symbiosis evolution and genetics with microbiome ecology (Tsiknia et al., 2020).

The model symbiotic mutualism between leguminous plants and the nitrogen fixing bacteria known as rhizobia is responsible for significant contributions to bioavailable nitrogen sources in the environment and in agricultural systems (Vitousek et al. 1997; Herridge et al. 2008b). In this interaction, legumes house rhizobia within nodules that develop on their roots and provide rhizobia with photosynthate, while in return rhizobia reduce atmospheric nitrogen N_2 to the fertilizer ammonium NH_4^+ (Gage,

2004). With the advent of next generation sequencing, it has become increasingly clear that rhizobia are not always the only endophytic occupants of root nodules. Nodules can be colonized by non-rhizobium endophytes (NREs), bacteria that come from dozens of genera and span multiple phyla (Martínez-Hidalgo et al. 2014; Martínez-Hidalgo and Hirsch 2017; Tokgöz et al. 2020; Rahal and Chekireb 2021). An enhanced understanding of how these NREs potentially impact the model legume-rhizobium mutualism would be useful in predicting the outcomes of legume-rhizobium evolution in more complex microbial communities, especially if both legume and rhizobium finesses are impacted by other microbiome constituents (Martínez-Hidalgo and Hirsch 2017; Burghardt and diCenzo 2023).

Of the NREs that have been described so far, co-inoculation has resulted in either positive or neutral effects on plant growth and nodule phenotypes (Khan, 2019; Martínez-Hidalgo & Hirsch, 2017). For example, *M. truncatula* co-inoculated with *Sinorhizobium medicae* WSM419 and *Pseudomonas fluorescens* WSM3457 formed a greater number of nodules (Fox et al. 2011). Also, *Medicago sativa* co-inoculated with *Sinorhizobium meliloti* 1021 and *Micromonospora* isolates had more efficient nutrient uptake and increased shoot mass (Martínez-Hidalgo et al., 2014). Thus, based on the limited number of studies to date, NREs have often been broadly described as plant-growth promoting bacteria (PGPBs) (Martínez-Hidalgo & Hirsch, 2017). However, the NREs tested thus far were chosen because of their plant-growth promoting potential on other non-legumes, meaning they are likely to be biased towards showing beneficial impacts. It remains unclear whether the diversity of NREs that naturally co-occur with rhizobia and plant

genotypes will similarly have positive effects on symbiotic outcomes, or whether some might have negative impacts.

Shared coevolutionary history and/or shared environmental conditions can heavily influence the effects of symbiosis on plant hosts. For example, populations of *Mesorhizobium* associated with *Acmispon wrangelianus* legume hosts displayed contrasting levels of heavy metal adaptation depending on shared soil type (Porter et al. 2017). Additionally, plant-associated microbial communities became more beneficial for their plant hosts under the appropriate drought regime (Lau and Lennon 2012; Bolin et al. 2023). In one study, legume genotypes inoculated with rhizobia that had been experimentally evolved on the same genotype across hundreds of generations grew better compared to those inoculated with rhizobia that evolved on different host genotypes (Batstone et al., 2020). NREs isolated from nodules with a co-occurring rhizobium strain could have positive impacts on the symbiosis if they increase the net benefits gained by both legumes and rhizobia, for example by providing rhizobia with a critical metabolite that enhances their ability to fix N. However, if NREs instead act as rhizobium competitors and/or plant pathogens, they might reduce the net benefits exchanged, and thus negatively impact the symbiosis. Predicting the ecologically relevant impacts that NREs have on the legume-rhizobium symbiosis will require a direct comparison of the net benefits gained by legumes and rhizobia in the presence or absence of multiple, distinct NREs that co-occupy nodules collected from hosts in native soils.

Here, we use four naturally occurring NREs and a high-quality strain of *Sinorhizobium meliloti* (strain 141; Batstone et al. 2022c) (formerly *Ensifer meliloti*), which were isolated from soils in the native range of the model legume *Medicago*

truncatula (Riley et al. 2023). Our objectives were to (1) obtain a taxonomic assignment of the four NREs; (2) conduct a greenhouse experiment in which fitness proxies of both plant (e.g., above-ground shoot biomass) and rhizobia (e.g., nodule number) were quantified and compared when plants were inoculated with *S. meliloti 141* alone versus plants co-inoculated with both *S. meliloti* and one of the four NREs; and finally, (3) verify whether NREs could be found in root and nodule tissues of plants when co-inoculated with *S. meliloti* 141.

METHODS

Study system

We chose the *Medicago truncatula* genotype DZA 315.6 (hereafter DZA) because we previously characterized its growth when paired with hundreds of rhizobial strains (Batstone et al. 2022a, 2022b), and it has a well-documented ability to form root nodules with diverse soil endophytes (Etienne-Pascal & Pascal, 2013). We chose *Sinorhizobium meliloti* strain 141 (hereafter *Sinorhizobium*) as our focal rhizobium partner because it was found to be a high-quality symbiont across several *Medicago truncatula* genotypes (A17 and DZA; Batstone et al. 2022b). The four non-rhizobial endophytes (NRE) used in this study, along with *Sinorhizobium*, were previously isolated from the root nodules of natural populations of *Medicago truncatula*, as described in detail in Riley et al. (2023). Briefly, strains were isolated from soils surrounding *M. truncatula* roots from 21 sites spanning the species' native range: Spain, France, and Corsica. All nodules were surface sterilized prior to crushing and streaking for strain isolation. To determine strain taxonomic identity, Riley et al. (2023) performed several rounds of purification, extracted

DNA from each purified culture, and submitted DNA for Illumina short-read shotgun sequencing. After trimming and quality filtering, Riley et al. (2023) conducted *de novo* assembly on the raw reads and submitted the resulting whole genome sequences to RASTtk for annotation (Brettin et al., 2015). While most strains were assigned to the expected *Sinorhizobium* spp., four were found to be "off-target" strains (522, 702A, 717A, and 733B), henceforth referred to as non-rhizobia endophytes (NREs). After determining that the four NREs were phylogenetically distant from rhizobia, based on colony morphology and a phylogeny constructed from several marker genes, Riley et al. (2023) excluded these strains from further analyses. All other data on these strains come from the present study.

Multi-locus species tree and average nucleotide identity

We first assigned our four NREs to putative genera by aligning the 16S rRNA gene sequences of each against the NCBI RefSeq 16S rRNA database (O'Leary et al., 2016), and inferring genera based on the top five BLAST hits. Because species-level identification of bacteria often requires a high-resolution phylogenetic analysis (Lan et al., 2016), we additionally constructed a maximum-likelihood phylogenetic tree based on whole-genome alignments using the AutoMLST online service (Alanjary et al., 2019). Our pipeline first involved submitting whole genome sequences of the four NREs and then constructing species trees using *de novo* mode, which automatically determines the nearest organisms and MLST genes (Table S1-S2) for alignment construction (Alanjary et al., 2019). Next, we filtered inconsistent MLST genes and generated a MAFFT FFT-NS-2 alignment (Katoh et al., 2002). AutoMLST uses IQ-TREE (Nguyen et al., 2015) to

select the best fit model and subsequently generate 1000 ultrafast bootstrap replicates (Minh et al., 2013). AutoMLST also estimates the average nucleotide identity (ANI) for comparisons of the concatenated multi-locus sequences between query taxa and taxa on the resulting tree. Given that AutoMLST does not automatically report the ANI between query sequences, we estimated each strains' ANI using OrthoANI (Lee et al., 2016) to compare the phylogenetic distance between the query NREs to each other.

Plant genotype and growth methods

We razor scarified and surface-sterilized seeds of *M. truncatula* genotype DZA, washing them in 95% ethanol for 30 seconds and then commercial bleach for seven minutes, followed by sterile water for four minutes to rinse off any excess bleach. Before planting, we packed the bottoms of SC10R Ray Leech "Cone-tainers" (Stuewe & Sons, Inc., Tangent, OR) with a small handful of autoclaved polyester fiber to cover their drainage holes. We filled each Cone-tainer with ~ 200 mL of an autoclave-sterilized mixture of one-part root wash: one-part sand: four-parts turface MVP calcined clay (Profile Products LLC, Buffalo Grove, IL, USA). The resulting potting media composition is suitable for *M. truncatula* growth because it facilitates root extraction and rinsing upon harvest, which is required for measuring nodule phenotypes (see below). We sowed seeds ¾ centimeter deep in the potting media-filled Cone-tainers using sterile forceps, and immediately watered to compact the soil and prevent seed desiccation. We sowed two seeds in each pot in case one failed to germinate, and then covered the potting media surface and seeds with ½ cm of autoclaved-sterilized vermiculite to help retain moisture

around seeds after sowing. Prior to inoculation, we thinned all pots to one seedling using sterile forceps.

Inocula preparation

For each of our five strains (four NREs and one *Sinorhizobium*), we separately steaked out stocks onto sterile petri dishes with solid Tryptone-Yeast (TY) medium (Vincent, 1970). Dishes were sealed with parafilm and placed into a 30 °C dark incubator. We allowed colonies to grow until ample growth was observed, approximately 48 hours. For each strain separately, we picked a single colony to inoculate a sterile 15 mL Falcon tube filled with liquid TY medium, and then placed the capped tubes in a shaking incubator set to 30 °C and 200 rpm for overnight growth. Between 20-24 hours later, we combined tubes of the same strain into a single sterile 50 mL falcon tube, gently inverted to mix, and pipetted out 500 µL to estimate cell density (cells per mL) via measuring absorbance at OD_{600} using a NanoDrop 2000c (Thermo Scientific; Waltham, MA, USA). To ensure that each strain started at an equal inoculation density, we added an appropriate amount of sterile liquid medium and culture to reach a final OD_{600} of 0.1, which corresponds to ~1×10^8 cells/mL.

Greenhouse experiment

We tested the direct and indirect effects of NREs on the legume-rhizobium symbiosis by directly comparing co-inoculations of *Sinorhizobium* paired with each of the four NREs to single-inoculations with *Sinorhizobium* alone or each of the four NREs alone, for a total of 10 treatments: five single-inoculation, four co-inoculation, plus one uninoculated

control to monitor contamination levels. The single-inoculation treatments consisted of either a 500 μL dose of *Sinorhizobium* (hereafter "*Sinorhizobium*-only") or a 500 μL dose of one of the four NREs (hereafter "NRE-only"), each dose totaling ~$5x10^7$ cells ($1x10^8$ cells/mL x 0.5 mL). The co-inoculation treatments consisted of one full 500 μL dose of one of the four NRE strains *and* a full 500 μL dose of *S. meliloti* 141 (hereafter "co-inoculation"), the combined dose totaling $1x10^8$ cells ($1x10^8$ cells/mL x 1 mL). Given that *Sinorhizobium* was the plant's only source of fixed nitrogen, we opted to control for the number of *Sinorhizobium* cells across inoculation treatments rather than the total number of cells; any difference in the number of *Sinorhizobium* cells present in the inoculum would necessarily change the fitness benefits plants receive from symbiosis regardless of whether NREs were present.

We randomized four racks containing five plants per treatment (and control) across two greenhouse benches, resulting in 40 racks total (10 treatments x 4 replicate racks x 5 replicate plants = 200 plants), plus one additional rack of ten external control plants placed outside of the benches in the same room (210 total plants). We programmed misters located inside the greenhouse room to go off for a duration of 45 minutes, four times a day. We inoculated plants with their respective treatments seven days after seeds had been sown (i.e., when the first true leaf had emerged) to give the plants sufficient time to establish their root systems and begin photosynthesizing. To further reduce cross-contamination, immediately after inoculation, we added ½ cm of sterile sand to the surface of the soil of each Cone-tainer to provide a barrier between the inoculum and air.

We destructively harvested all plants four weeks post inoculation. We cut the above-ground shoots of all plants at their bases and separately stored them in five-inch

coin envelopes, which were left to dry in a 60 °C oven for three days before the dry-mass of each shoot was recorded. We extracted roots from Cone-tainers, dunked them into a bin of tap water to gently remove potting media, and then wrapped roots in paper towels and stored them in a 4 °C refrigerator until nodule dissection the following day. After counting the total number of nodules formed on each root, we haphazardly selected ten nodules, removing them carefully with forceps, and weighing all ten together to the nearest 0.01 mg using a microbalance (Mettler-Toledo, Columbus OH, USA). After nodule dissection, we placed roots into five-inch coin envelopes, allowed them to dry using the same protocol for shoots, and recorded their weight.

Nitrogen addition experiment

For NREs that were found to significantly impact plant traits, we wanted to tease apart whether these effects were acting on the plants directly (i.e., observed in the absence of co-inoculation with rhizobia) or indirectly (i.e., via their interactions with rhizobia). However, because we did not supply plants with any external sources of nitrogen in our greenhouse experiment (described above), plants grew poorly in all treatments in which *Sinorhizobium* was absent, precluding our ability to thoroughly test direct vs indirect effects of NREs. To address this limitation, we conducted an additional experiment in which plants were supplied with moderate amounts of N and were either inoculated with a single NRE or were left uninoculated (control). If plants grew larger or smaller when inoculated with an NRE compared to uninoculated controls, then we would consider this NRE to incur direct benefits (i.e., mutualistic) or costs (i.e., pathogenic) on the plant respectively. We prepared 16 replicate plants per inoculation treatment, using similar

protocols as described for the greenhouse experiment (above). After scarification and surface-sterilization, we placed seeds on sterile petri dishes with sterile water and allowed them to germinate for 48 hours in the dark. Seedlings were then transplanted into magenta boxes (PlantMedia, Dublin, OH) containing autoclaved-sterilized 1:1 calcined clay and sand mixture to exclude any plant-available nitrogen. Plants were placed in a growth chamber at a constant 60% humidity with 16-hour, 23 °C days and 8-hour, 18 °C nights. One week after transplanting, plants were inoculated with each NRE individually or with sterile media. Plants received 5 mL of N-supplemented Fahräeus media (Vincent, 1970) once a week for four weeks after inoculation. We used ammonium nitrate (NH_4NO_3) as the external N source, ramping up the concentration at each application to reflect varying nitrogen requirements during seedling development (Barker et al., 2006): 1.0 mg/mL, 2.0 mg/mL, 3.0 mg/mL, and 5.0 mg/mL, week-by-week. After these four weeks, we destructively harvested plants, recorded their dry shoot masses (as described above), and examined roots for nodules to check for contamination.

Statistical analyses on phenotypic data

For data measured in the greenhouse experiment, we constructed a linear mixed-effects model (LMM) using the package "lme4" (Bates et al., 2015) in R (R Core Team, 2020) to test if observed traits were influenced by inoculation treatment. For each model, we included treatment as a fixed effect and plant rack as a random effect to account for the spatial arranged of racks in the greenhouse. We performed type II ANOVA on the model generated for each trait using the “car” package (J. Fox & Weisberg, 2019). To infer significance in pairwise comparisons of each treatment for each trait, we estimated

marginal means with false discovery rate adjusted p-values using the “emmeans” package (Lenth, 2022). For the nitrogen experiment, we analyzed data using LMMs and post-hoc analysis as described above, but with inoculation treatment (control, inoculated) as fixed effect and pot position as a random effect.

Tissue occupancy experiment

Finally, we wanted to confirm whether NREs that significantly impacted the symbiosis, which were originally isolated from nodules that formed on plants growing in the field, could reinfect plant tissues after inoculation. We conducted an additional co-inoculation experiment of *Sinorhizobium* paired with each NRE that affected host growth (see Results). Seeds of DZA *M. truncatula* were prepared and grown with the same methods as the nitrogen addition experiment (see above). We grew ten plants of each treatment in pre-sterilized plastic, drainless, three-inch plant trays that were covered with transparent, plastic 10-inch domes, yielding 50 plants in total. We filled plant trays with sterile tap water, so that plants would receive a consistent supply of water. We used the same inoculation methods as those described above.

Plants were harvested four weeks post-inoculation with the same protocol as described above, but after rinsing roots, we placed them in a sterile 50 mL falcon tube with 40 mL sterile DI water and shook thoroughly to remove excess soil and weakly-associated rhizosphere microbes. After washing roots this way, we placed each root into its own sterile 50 mL falcon tube with silica beads and cotton at the bottom (to remove excess moisture) and left to dry overnight. We removed nodules from roots as well as root tissue sections without nodules. Nodules and roots were placed in separate 1.5 mL

tubes with 500 μL sterile water and allowed to re-imbibe overnight. We then partitioned these tissue samples for DNA sequencing into two categories: endophytes (within tissues) as well as endophytes + epiphytes (on tissue surfaces). In total, we gathered 48 tissue samples for DNA sequencing (3 treatments x 2 tissue sections x 2 sterilization treatments x 4 replicates = 48 samples). To sequence only endophytic bacteria, we randomly chose four samples from each tissue section per inoculum treatment and surface sterilized each individually by adding 1 mL of 30% commercial bleach, thoroughly mixing for 60 seconds, and then removing bleach thoroughly by washing each sample with 1 mL of sterile DI water. To sequence epiphytic in addition to endophytic bacteria, we randomly chose four additional samples per tissue section per treatment, washing each individually with 1 mL of sterile DI water, and thoroughly mixing for 60 seconds to remove loosely-associated bacteria on tissue surfaces.

DNA extraction and 16S V3-V4 amplicon sequencing

We extracted and purified genomic DNA from tissue samples using a DNeasy Plant Pro Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. For each sample, we evaluated the DNA concentration using optical density measurements obtained by a Nanodrop spectrophotometer (NanoDrop Technologies, Rockland, DE, USA) at 260 and 280 nm wavelengths. We submitted DNA samples to the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign for 16S rRNA V3-V4 region amplicon sequencing. Briefly, DNA quality was assessed with a Qubit fluorometer (Invitrogen, MA, USA) and the targeted 16S regions were amplified by PCR using primers V3-F357_N (5'-CCTACGGGNGGCWGCAG-3′)

and V4-R805 (5'-GACTACHVGGGTATCTAATCC-3′) using a Fluidigm Biomark HD PCR machine (Fluidigm Corporation, South San Francisco, CA, USA). PCR products were quantified with a QuantiT PicoGreen fluorometer (Invitrogen, MA, USA), then cleaned, purified, and size-selected in a 2% agarose E-gel (Invitrogen, MA, USA) followed by gel extraction with a QIAquick Gel extraction kit (Qiagen, Hilden, Germany). Amplicons were sequenced with Illumina MiSeq v2 platform (Illumina, San Diego, CA, USA) according to the manufacturer's instructions.

Amplicon sequence variant (ASV) inference

For each library (i.e., sample), we inspected raw, demultiplexed reads for quality using DADA2 v1.28.0 (Callahan et al. 2016) in R (R Core Team 2020). Reads were trimmed with DADA2 to remove forward and reverse primers at lengths of 32 and 35, respectively, to account for the length of V3-V4 primers plus the length of Fluidigm-specific CS primers. We filtered the resulting trimmed reads with DADA2 to retain reads with a maximum number of expected errors of two on the forward strand and five on the reverse strand, in addition to removing putative phiX sequences. We truncated reads at the first instance of a quality score less than or equal to two. The resulting filtered and trimmed reads were dereplicated with DADA2 and amplicon sequence variants (ASVs) were inferred from the dereplicated reads with DADA2. We merged ASVs from forward and reverse read pairs and removed putative chimeras with DADA2, using the "consensus" method. The resulting merged, non-chimeric ASVs were used to construct a sequence table, and the taxonomy of the ASVs was inferred with DECIPHER v2.28.0 (Wright 2016) with the SILVA SSU r138 database (Quast et al. 2012). We manually

renamed the resulting taxonomic assignments to the genus *Ensifer* to *Sinorhizobium* to reflect the most recent modifications to the genera *Ensifer* and *Sinorhizobium* (Kuzmanović et al. 2022). Additionally, we removed ASVs without a taxonomic assignment at the domain level and ASVs assigned to the orders Rickettsiales or "Chloroplast" using phyloseq v1.44.0 (McMurdie and Holmes 2013) to exclude ASVs inferred from host DNA from downstream analyses. We constructed a heatmap visualizing the matrix of variance-stabilized ASVs counts per sample using the R package pheatmap v1.0.12 (Kolde 2019) using color palettes generated by RColorBrewer v1.1.3 (Neuwirth 2022) and viridis v0.6.3 (Garnier et al. 2023). The resulting heatmap was manually edited to position color legends and to format the font face of ASV labels for Figure 4. ASV names with polyphyletic or *Candidatus* genus assignments from SILVA were manually replaced with "NA".

ASV differential abundance analysis

To determine the relative enrichment of ASVs across inoculum treatment, tissue section, and surface-sterilization treatment, we used the absolute, non-normalized ASVs counts in conjunction with DESeq2 (Love et al. 2014) v1.40.2 for a differential abundance analysis. Briefly, we built negative binomial generalized models using the ASV counts in DESeq2 with inoculum, tissue section, and surface-sterilization treatment included as factors in the models. To determine ASVs whose abundances were significantly affected by each factor individually, we performed likelihood ratio tests using DESeq2 using full models of the three factors together and reduced models lacking one of each factor. For each test, we shrunk the resulting log2 fold-changes of every contrast of the factor's

levels using the DESeq2 function “lfcShrink”. Lastly, to determine ASVs differentially abundant across surface-sterilization treatments among the same tissue sections, we performed an additional likelihood ratio test. The full model in this test contained one factor representing the interaction of tissue section and surface-sterilization treatment, while the reduced model included only the intercept. As done with the other tests, we shrunk the log2 fold changes of each level contrast. We generated all plots with ggplot2 (Wickham 2016).

ASV BLAST alignments

Using NCBI BLAST+ v2.14.0, we created a nucleotide database from the 16S rRNA sequences annotated from the genome assemblies of *Sinorhizobium*, *Paenibacillus* sp. 717A, and *Pseudomonas* sp. 733B. All ASV nucleotide sequences inferred by DADA2 (Table S9) were used as a query in a BLASTn search with BLAST+ (Camacho et al. 2009), retaining alignments with a minimum e-value of 0.01 and at least 80% identity. ASV28 (Table S9) was determined to be a likely representative of *Paenibacillus* sp. 717A (see results) and was aligned against the NCBI rRNA type strain database on the BLAST+ online platform (blast.ncbi.nlm.nih.gov) to verify its potential taxonomic assignment. We determined that ASV1 and ASV5 were representatives of *Sinorhizobium* 141 and *Pseudomonas* sp. 733B, respectively (see results). To investigate the presence of strains related to these two ASVs plus ASV28 in other contexts, we aligned ASV1, ASV5, and ASV28 against the 50 most abundant sequences inferred from 16S rRNA amplicon sequencing in another study of symbiotic bacteria of *M. truncatula* (Brown et al. 2020) using BLASTn with the same parameters.

RESULTS

Phylogeny

Our 16S rRNA BLAST results suggested both NRE strains 522 and 717A belong to the genus *Paenibacillus*, while 702A and 733B belong to *Pseudomonas*. We subjected both *Pseudomonas* strains and both *Paenibacillus* strains to multi-locus species tree construction using whole-genome sequences. This process revealed that the whole-genome sequence for *Paenibacillus sp.* 522 was contaminated with contigs from a *Brevibacillus*, and so only *Paenibacillus* reference genomes were included in the ingroup for tree construction to avoid errors; however, it is possible that this *Brevibacillus* was present in the 522 inoculations. The resulting *Pseudomonas* tree (Figure 1A) placed *Pseudomonas sp.* 733B as a sister taxon to *Pseudomonas reinekei* MT1 (mean distance = 0.0894, ANI = 0.9106, $P < 0.001$) with 100 percent bootstrap support at this node. *Pseudomonas sp.* 702A was placed with 100 percent bootstrap support at the node sister to a clade containing five different *Pseudomonas* type strains, of which the closest neighbor was *Pseudomonas migulae* NBRC 103157 (distance = 0.1105, ANI = 0.8895, $P < 0.001$). The resulting *Paenibacillus* tree (Figure 1B) placed *Paenibacillus spp.* 522 and 717A sister to each other in the same clade with no other taxa with 100 percent bootstrap support at this node. The nearest type strains to *Paenibacillus spp.* 522 and 717A placed on the species trees were *Paenibacillus ferrarius* CY1 (distance = 0.2057, ANI = 0.7943, $P. < 0.001$) and *Paenibacillus alginolyticus* DSM 5050 (distance = 0.1849, ANI = 0.8151, $P < 0.001$), respectively. For a comprehensive list of calculated mash distances, see Tables S3-S4. The estimated OrthoANI for *Paenibacillus spp.* 522-717A and

Pseudomonas spp. 702A-733B comparisons were 0.8472 and 0.8524, respectively (Table S5).

Greenhouse experiment

Plants inoculated with *Sinorhizobium* plus either *Paenibacillus sp.* 522 or *Pseudomonas sp.* 702A were not significantly different from *Sinorhizobium*-only plants, whereas plants inoculated with *Sinorhizobium* plus either *Paenibacillus sp.* 717A or *Pseudomonas sp.* 733B had significantly smaller shoot mass compared to *Sinorhizobium*-only plants (Figure 2A; Table S7). Every group of plants singly inoculated with an NRE without N-fixing *Sinorhizobium*, and uninoculated control plants, were in observably poor condition upon harvest, with significantly smaller shoot masses compared to co-inoculated plants and *Sinorhizobium*-only plants (Table S7), while the differences in shoot mass across the NRE-only and uninoculated control groups were not significant (Figure2A; Table S7). Uninoculated plants showed negligible signs of contamination; we only found two nodules on a single internal control plant, while all other controls (n = 29) were nodule-free. Similarly, we only found two plants inoculated with an NRE but not *Sinorhizobium* that formed nodules (one with 46 and the other with 31 nodules), which were subsequently removed from downstream analyses. Additionally, the biomass of all plants that were not inoculated with *Sinorhizobium* were significantly smaller than those inoculated with *Sinorhizobium* (Table S7).

The lack of nodules on the vast majority of plants (n = 78/80) that had been inoculated only with NREs (Figure 2B) indicated that these NREs were indeed unable to form nodules on their own. The difference between the average number of nodules

formed by *Sinorhizobium*-only and *Paenibacillus sp.* 522 or *Pseudomonas sp.* 702A co-inoculated plants was not significant (Figure 2B; Table S7). However, both *Paenibacillus sp.* 717A and *Pseudomonas sp.* 733B co-inoculated plants formed significantly fewer nodules on average compared to *Sinorhizobium*-only plants with a 23% and 25% reduction (n = 20), respectively, while none of the co-inoculated groups were significantly different from each other (Figure 2B; Table S7). This pattern is consistent with the differences observed in shoot mass, above. For a comprehensive list of all pairwise contrasts for each trait measurement and their associated *P*-values, see Table S7.

Nitrogen addition experiment

While co-inoculations of *Sinorhizobium* with *Paenibacillus sp.* 717A or *Pseudomonas sp.* 733B negatively impacted host-symbiont traits (Figure 2B; Table S7), the effects on *M. truncatula* shoot mass could have been either an indirect result of the antagonistic interaction between NRE and *Sinorhizobium* or a direct result of parasitism by the NREs on the *M. truncatula* host. Since plants in the greenhouse experiment did not receive supplemental nitrogen, we could not distinguish whether the poor performance of *M. truncatula* plants singly inoculated with NREs was due to nitrogen-deficiency alone or if NRE inoculations also contributed. To address this, we tested whether the NREs were plant pathogens by examining the direct effects of *Paenibacillus sp.* 717A and *Pseudomonas sp.* 733B on host plants that received supplemental nitrogen fertilizer (Figure 3). Plant growth was not significantly impacted by inoculation with either strain compared to the uninoculated controls ($P > 0.1$, type II ANOVA), suggesting that the NREs 717A and 733B did not directly affect *M. truncatula*; rather, they likely indirectly

affected plant fitness in co-inoculations by inhibiting the legume-rhizobium symbiosis. Again plants grown without supplemental fertilizer (and without *Sinorhizobium*) were in poor condition (Table S8), and no nodules were observed on any plant.

Tissue occupancy experiment

Finally, we used 16S rRNA amplicon sequencing of both surface-sterilized and non-surface-sterilized root and nodule tissues from coinoculated plants to locate these NREs in the endosphere and/or rhizosphere. Although we inoculated plants with combinations of only *Sinorhizobium*, *Paenibacillus sp.* 717A, and *Pseudomonas* sp. 733B in initially axenic conditions, taxonomic assignments of ASVs after four weeks in the greenhouse showed that other bacteria were also associated with our tissue samples (Figure 4). A total of 248 unique ASVs were inferred across all samples (Table S9); however, the vast majority of these ASVs were in very low abundance (Table S9). This low abundance is expected to some degree considering that 16S rRNA amplicon sequencing can report ASVs with abundances < 0.1% (Nikodemova et al. 2023), which either reflect spurious inferences or trace amounts of microbial cells. Nonetheless, ASVs were inferred from all inoculation, tissue, and surface-sterilization groups (Figure 4).

Our goal was to determine whether we could recover our inoculum strains from nodules in the expected treatments, which would confirm these strains as nodule endophytes rather than tissue surface contaminants. After aligning our inferred ASVs against the 16S rRNA sequences of our isolate genomes, we found that ASV1 and ASV5 aligned perfectly to the *Sinorhizobium* and *Pseudomonas* sp. 733B 16S V3-V4 regions, respectively, retaining 100% of the ASV lengths in the alignments (Table S11). Our

taxonomic assignments of ASVs based on alignments to SILVA are consistent with these predictions (Figure 4). ASV1 (*Sinorhizobium* based on SILVA references) was abundant in every inoculum treatment, tissue section, and surface-sterilization treatment. The abundance of ASV1 was not different between samples inoculated with *Sinorhizobium* alone and either 717A + 141 or 733B + 141 samples ($P > 0.05$, Wald test, FDR adjusted) (Figure 4, Figure S1A-B, Table S14). As expected, ASV1 was significantly more abundant among nodule samples compared to root samples ($P < 0.01$, Wald test, FDR adjusted) (Figure S1C). These results suggest that neither *Paenibacillus* sp. 717A nor *Pseudomonas* sp. 733B significantly reduced the abundance of *Sinorhizobium* in nodules. ASV5 was assigned to *Pseudomonas* SILVA reference sequences and was only detected in samples from *Pseudomonas* sp. 733B co-inoculated plants. This ASV was present in non-surface-sterilized nodule samples plus one additional surface-sterilized root sample at trace levels but was significantly more abundant in nodule samples compared to root samples ($P < 0.01$, Wald test, FDR adjusted) (Figure 4, Figure S1C, Table S14).

No ASVs aligned perfectly to the *Paenibacillus* sp. 717A 16S gene (Table S11). Among the ASVs that did align to *Paenibacillus* sp. 717A, the most abundant ASV in our dataset was ASV28 (Table S11; Table S9; Figure 4), which aligned at an 85% identity with 99% coverage of the ASV (Table S11). Several *Paenibacillus* species have displayed relatively high levels of intraspecific sequence heterogeneity in their 16S rRNA genes (Nübel et al. 1996; Bosshard et al. 2002; Nechayeva et al. 2021), which likely explains this smaller alignment identity compared to our other inoculum strains. This sequence was assigned to the genus *Bacillus* based on the SILVA database (Figure 4); however, this is likely attributed to the under-representation of *Paenibacillus* isolates

relative to *Bacillus* in the SILVA SSU training database used for sequence classification. We verified the assignment of ASV28 to the genus *Paenibacillus* by BLASTing its sequence against the NCBI 16S rRNA type strain database and indeed found the highest scoring alignments to members of *Paenibacillus*. Overall, out of all other ASVs inferred from the amplicon data, ASV28 is the most likely representative of *Paenibacillus* sp. 717A. This ASV was found only in non-surface-sterilized nodule and non-surface-sterilized root samples from *Paenibacillus* sp. 717A co-inoculated plants, with no difference in abundance between the two tissue sections ($P > 0.05$, Wald test, FDR adjusted) (Figure 4, Figure S1A, Table S14).

To determine whether inoculum strains were differentially abundant across the *M. truncatula* rhizosphere versus the endosphere, we compared surface-sterilized to non-surface-sterilized samples (Figure S1D-E). Strains that are endosphere-associated should either be enriched in the surface-sterilized groups (indicating preference for the endosphere over the rhizosphere) or not differentially abundant between the two groups (indicating similar levels across the endosphere and rhizosphere). However, strains that are rhizosphere-associated should be enriched in non-surface-sterilized groups. In the nodule samples, ASV5 [*Pseudomonas* sp. 733B] and ASV28 [*Paenibacillus* sp. 717A] were enriched in non-surface-sterilized samples ($P < 0.05$, $P < 0.01$, Wald test, FDR adjusted), while ASV1 [*Sinorhizobium*] was not differentially abundant ($P > 0.05$, Wald test, FDR adjusted) (Figure S1E, Table S14). In the root samples, ASV28 was enriched in non-surface-sterilized samples ($P < 0.01$, Wald test, FDR adjusted). Although ASV1 and ASV5 each had a significant difference in abundance between surface-sterilized root samples compared to non-surface-sterilized root samples ($P < 0.05$, Wald test, FDR

adjusted), neither ASV met our minimum fold-change criterion of 1.5 to be considered enriched in either surface-sterilized or non-surface-sterilized root samples (Figure S1F, Table S14). Collectively, these comparisons show that *Sinorhizobium* was indeed endosphere-associated in both root and nodule sections. *Paenibacillus* sp. 717A and *Pseudomonas* sp. 733B, on the other hand, were rhizosphere associated at nodules, with *Paenibacillus* sp. 717A also being rhizosphere associated in root sections and *Pseudomonas* sp. 733B being endosphere-associated in root sections.

Sinorhizobium, Paenibacillus, and Pseudomonas associations beyond this study

To further investigate the ecological relevance of our NREs, we asked whether our three ASVs of interest, representing the three inoculum strains, could also be detected among amplicon data from a separate study of the leaf, root, and nodule endophytes of *M. truncatula* endophytes grown in native field soil (Brown et al. 2020). Aligning our ASVs of interest against the 50 most abundant sequences generated by Brown et al. (2020) revealed imperfect but numerous alignments to taxa in their study (Table S12). The best alignment of ASV1 was OTU00001, an "*Ensifer*" bacterium (now reclassified as *Sinorhizobium*) that was found to be a significant nodule endophyte in the Brown et al. (2020) study, at 99% identity with 92% coverage of the OTU (Table S12). Our ASV28 [*Paenibacillus* sp. 717A] aligned to OTU00009 at 100% identity and 92% coverage (Table S12). This OTU was a significant rhizosphere member in Brown et al. (2020) and was labelled as a "*Paenisporosarcina*" in the order Bacillales, inferred by aligning OTUs against a SILVA 16S reference alignment (v123). Lastly, our ASV5 [*Pseudomonas* sp. 733B] aligned equally well to OTU00013 and OTU00003, which were labeled by Brown

et al. (2020) as "*Streptomyces*" and "*Pseudomonas*", respectively, at 99% identity and 92% coverage (Table S12). Considering OTU sequence lengths in Brown et al. (2020) were over 100 bases shorter than the ASVs generated here, the alignments of ASV5 to both *Streptomyces* and *Pseudomonas* OTUs are likely due to the lower sequence resolution in their study. Given ASV5 [*Pseudomonas* sp. 733B] aligned perfectly to a region of the 16S rRNA gene in our *Pseudomonas* sp. 733B genome assembly, we believe that our ASV5 [*Pseudomonas* sp. 733B] is likely closely related to *Pseudomonas* OTU00003, a significant member of the leaf phyllosphere in Brown et al. (2020).

DISCUSSION

Considering the ecological and economic importance of the legume-rhizobium symbiosis (Sprent 1987; Herridge et al. 2008a), it is critical to understand how the outcomes of this symbiosis change due to both abiotic and biotic factors, including non-rhizobial endophytes (NREs) that share the nodule alongside rhizobia. Using four NRE strains, we inferred their phylogeny, re-identified the NREs from root and/or nodule tissues, and showed that two of the four NREs we examined reduced the benefits both legumes and rhizobia receive from the interaction. We discuss the potential mechanisms for, and implications of, these main results below.

Several studies of legume-rhizobium symbiosis have demonstrated how genetic variation in either of these partners can influence fitness outcomes, in addition to interactions with other plants and microbes (Heath and Tiffin 2009; Brown et al. 2020; Batstone et al. 2022a). For example, one study found that variation in host control among six lines of *Acmispon strigosus* influenced the variation in symbiont effectiveness among

Bradyrhizobium populations (Wendlandt et al. 2019). Far fewer studies have explicitly considered genetic variation in other microbial constituents present in the soil community (Tsiknia et al. 2020). Our results suggest that genetic variation within NREs (*Pseudomonas spp.* 702A vs. 733B and *Paenibacillus spp.* 522 vs. 717A) leads to variation in fitness-related traits of both the *M. truncatula* host and the *S. meliloti* symbiont. While it is infeasible to capture the effects of all surrounding taxa on legume-rhizobium symbiosis, investigating interactions with microbes present inside root nodules—the location where the critical symbiotic exchange of nutrients is occurring—is an important step towards understanding legume-rhizobium symbiosis in its full context. We show that NREs present in nodules alongside rhizobia can have indirect effects that must be considered if the outcomes of the legume-rhizobium symbiosis in real soil communities are to be predicted more accurately.

Effects on plant fitness

Symbionts in a microbiome do not always evolve to benefit their plant host, as microbes have their own fitness interest that do not necessarily align with host fitness (Klein et al. 2021). In contrast to previous studies, the NREs included here were isolated from the same region as our focal *Sinorhizobium* strain, and negatively impacted the legume-rhizobium symbiosis by decreasing the average number of nodules formed and average shoot mass. The vast majority of NREs described to date have demonstrated positive or neutral effects on plant or rhizobium performance by increasing host shoot mass or the number of nodules formed (Khan, 2019; Martínez-Hidalgo & Hirsch, 2017). Since two NREs in this study (*Paenibacillus sp.* 717A and *Pseudomonas sp.* 733B) negatively

impacted the legume-rhizobia mutualism while the other two (*Paenibacillus sp.* 522 and *Pseudomonas sp.* 702A) showed neutral effects, NREs should not generally be considered beneficial.

The negative effects of *Paenibacillus sp.* 717A and *Pseudomonas sp.* 733B on plant growth could have been due to several non-mutually exclusive mechanisms, including that these strains: i) represent plant pathogens, whereby plants perform worse in the presence of these NREs compared to when they are not associated with them regardless of the presence of rhizobia; ii) reduce *Sinorhizobium*'s ability to colonize nodules once formed, and subsequently, fix N; and/or iii) reduce *Sinorhizobium*'s ability to nodulate plants. The results from our nitrogen addition experiment, whereby plants were provided with sufficient levels of nitrogen critical for minimal growth (Küster 2013), showed little evidence for i), that NREs are plant pathogens; plants performed similarly when inoculated by NREs compared to sterile media. We also found little evidence to suggest ii), that NREs impede rhizobia's ability to colonize nodules because our tissue occupancy experiment revealed that *Sinorhizobium* ASV abundance within nodules was not significantly impacted by NRE co-inoculation. Instead, our results are most consistent with iii), that NREs interfere with *Sinorhizobium*'s ability to nodulate plants, given the significant reduction in nodules formed by plants when co-inoculated with NREs compared to *Sinorhizobium* alone.

Finding a reduction in nodulation despite plants being N starved in our greenhouse experiment suggests that these NREs inhibited nodulation, which precludes N acquisition by the host from rhizobial partners. When the only source of N is that fixed by rhizobia, legumes will continue to form nodules until they acquire sufficient levels of

N via a tightly regulated process known as autoregulation of nodulation (AON, Bauer 1981; Caetano-Anollés and Gresshoff 1991). Fewer nodules and poorer plant growth means the reduction in nodules we observed was unlikely due to AON decreasing nodulation. A recent study in which legumes were either co-inoculated with two strains of rhizobia or singly inoculated with each strain individually found a similar reduction in nodulation due to interference competition between strains, whereby the growth of each strain is inhibited by the presence of the other (Rahman et al. 2023). Given their ability to co-inhabit root and nodule tissues, direct interactions between rhizobia and NREs are likely. In other studies, rhizobia-NRE interactions were found to impact nodulation in a variety of ways. For example, *S. meliloti* was found to cross-utilize siderophores produced by *Exiguobacterium* NREs in co-inoculations, which stimulated nodule formation and increased the mass of their *Trigonella foenum-graecum* host (Rajendran et al. 2012). Additionally, production of indole-3-acetic acid by *Pseudomonas trivialis* 3Re27 and *Pseudomonas extremorientalis* TSAU20 stimulated nodule formation by *Rhizobium galegae* on roots of *Galega orientalis* and increased *G. orientalis* growth (Egamberdieva et al. 2010). Furthermore, *Bacillus subtilis* UD1022 was found to antagonistically downregulate quorum sensing and biofilm formation by *Sinorhizobium meliloti* Rm8530, which is critical to initiate root nodule formation (Rosier et al. 2021). Future studies could determine whether the *Pseudomonas sp.* 733B and *Paenibacillus sp.* 717A NREs here act antagonistically against *Sinorhizobium* via extracellular secretions that inhibit nodule formation, or whether they physically compete with *Sinorhizobium* in root tissues. Either or both of these mechanisms could explain the indirect impacts of these NREs on plant growth. Overall, our results highlight that NREs have the potential

to negatively impact plants *indirectly* through their interactions with rhizobia, which in turn, impacts the fitness benefits received by both legumes and rhizobia. Improving symbiotic outcomes in natural or managed fields will therefore require examining the community context in which the symbiosis takes rather than the focal partners alone.

Cohabitation of Sinorhizobium and NREs

Evolutionary theory predicts that mutualisms are susceptible to "cheaters" that reap rewards from their symbionts without paying any costs (Jones et al. 2015). One current objective of coevolutionary research is to understand how genetic variation and abiotic factors stabilize or destabilize mutualisms (Sachs et al. 2004; Heath and Tiffin 2009; Jones et al. 2015; Batstone et al. 2018, 2022b). Some perspectives suggest that maintaining variation in partner quality, even continuing to associate with potential cheaters, offers a selective advantage to the host if the benefits obtained depend on environmental conditions (Batstone et al. 2018). By extension, the benefits exchanged in symbiosis are likely to depend on the biotic context, and more specifically, the microbiome in which the focal symbiosis unfolds. The natural co-habitation of rhizobia and NREs that impact legume-rhizobia symbiosis here could provide yet another mechanism for maintaining variation in partner quality, if quality depends on the identities of the NREs present within a microbiome (Batstone et al. 2018).

Soil bacteria interact with each other and influence the fitness of their host plant with and without infection (de la Fuente Cantó et al. 2020). Whether the microbes behind these interactions inhabit the nodule/root endosphere or rhizosphere has different implications for community composition and the phenotypic outcomes of symbioses

(Brown et al. 2020). While the NRE strains used here were originally isolated from cultures of the *M. truncatula* nodule endosphere, we were not able to confirm reinfection of nodule endospheres by either NRE strain, although we did detect them on surface nodule and root tissues and *Pseudomonas* sp. 733B in the root endosphere at relatively low levels. Our findings do not rule-out that these NREs occupied nodules alongside *Sinorhizobium*. Low-abundance bacteria can have significant impacts on the surrounding community (Lynch and Neufeld 2015; Jousset et al. 2017). These community members may be difficult to detect in amplicon sequencing data if their sequence counts in a sample are too low compared to other community members (Huse et al. 2010; Paulson et al. 2013; Escudié et al. 2018). Given *Sinorhizobium*'s extensive population growth and genome duplication within nodules (Mergaert et al. 2006), sequencing technologies that enrich for particular targets, such as capture sequencing (Hayden et al. 2022), may be required to detect non-rhizobial low frequency nodule occupants.

Identifying potential Pesudomonas and Paenibacillus NREs elsewhere

The *Pesudomonas* and *Paenibacillus* NRE strains studied here were originally isolated from nodules of *M. truncatula* grown in soils collected from its natural range. By comparing our sequencing results to those of another study that grew *M. truncatula* in such soils (Brown et al. 2020), we found OTUs representing taxa closely related to *Paenibacillus* sp. 717A and *Pseudomonas* sp. 733B to be among the most abundant community members. This relationship suggests that NREs in this study are likely reflective of NREs in the natural microbiome of *M. truncatula*. This presents opportunities to understand the mechanisms by which complex root-associated

communities establish and influence each other using this as a model multi-player symbiosis.

Concluding remarks

Here we show that some NREs can reduce the benefits of legume-rhizobium symbiosis. In our study, *Paenibacillus spp.* 522 and 717A as well as *Pseudomonas spp.* 702A and 733B are each phylogenetically distinct strains representing four novel species without high nucleotide similarity to existing RefSeq representatives. When co-inoculated alongside *S. meliloti*, two of these strains (717A and 733B) were re-identified on root and nodule tissues of *M. truncatula* and were found to inhibit nodulation and plant growth, likely via direct competitive interactions with rhizobia in the soil. Our results highlight that intraspecific variation within NREs can generate variable fitness outcomes for both partners in the legume-rhizobium symbiosis, meaning that predicting the impact of these "off-target" strains have on the symbiosis will require more than just testing a single representative strain across different species. Thankfully, with the rise of more affordable sequencing technologies, the variation in off-target microbes such as the NREs identified here can be uncovered more feasibly. Coupling sequencing with experiments to test the potential indirect effects of NREs present in natural or managed soil microbiomes will facilitate a better understanding of the coevolutionary dynamics in complex microbiome communities.

ACKNOWLEDGMENTS

The authors thank all current and past members of the Heath Lab at the University of Illinois at Urbana-Champaign who have assisted with and provided intellectual support for the work presented here. In particular, the authors thank Michael Grillo and Alex Riley for their initial efforts in isolating and identifying the bacterial strains used in this study. The authors also thank Karla Griesbaum, Nevers Mushimata, and Riley Popp for assisting with greenhouse preparations, plant harvests, and nodule dissections. The authors additionally thank Mario Cerón Romero for advice on bioinformatics analyses as well as Ivan Sosa Marquez for guidance with DNA extractions.

COMPETING INTERESTS

The authors declare there are no competing interests.

AUTHOR CONTRIBUTION STATEMENT

Conceptualization, J.C.K., R.B.D., and K.D.H.; Methodology, J.C.K., R.B.D., and K.D.H.; Software, J.C.K. and R.B.D.; Validation, J.C.K., R.B.D., and K.D.H.; Formal Analysis, J.C.K. and R.B.D.; Investigation, J.C.K., R.B.D., and K.D.H.; Resources, K.D.H.; Data curation, J.C.K., R.B.D., and K.D.H.; Writing — Original Draft, J.C.K.; Writing — Review & Editing, J.C.K., R.B.D., and K.D.H.; Visualization, J.C.K. and R.B.D.; Supervision, R.B.D. and K.D.H.; Project Administration, K.D.H. and R.B.D.; Funding Acquisition, K.D.H.

FUNDING STATEMENT

This work was supported by NSF PGRP-1856744, and DBI-2022049. Support for J.C.K. was provided by the 2020 Carl R. Woese Institute for Genomic Biology Undergraduate Research Fellowship, the University of Illinois at Urbana-Champaign School of Integrative Biology 2020 Spyros Kavouras Summer Research Award, and the University of Illinois at Urbana-Champaign School of Integrative Biology 2021 Robert H. Davis Excellence Award.

DATA AVAILABILITY STATEMENT

All scripts and intermediate files are available at github.com/jamesck2/coinoculation-endophytes.

REFERENCES

Afkhami, M.E., Almeida, B.K., Hernandez, D.J., Kiesewetter, K.N., and Revillini, D.P. 2020. Tripartite mutualisms as models for understanding plant–microbial interactions. Curr Opin Plant Biol **56**: 28–36. doi:10.1016/j.pbi.2020.02.003.

Bakker, M.G., Schlatter, D.C., Otto-Hanson, L., and Kinkel, L.L. 2014. Diffuse symbioses: roles of plant–plant, plant–microbe and microbe–microbe interactions in structuring the soil microbiome. Mol Ecol **23**(6): 1571–1583. doi:10.1111/mec.12571.

Batstone, R.T., Burghardt, L.T., and Heath, K.D. 2022a. Phenotypic and genomic signatures of interspecies cooperation and conflict in naturally occurring isolates of a model plant symbiont. Proceedings of the Royal Society B: Biological Sciences **289**(1978). Royal Society Publishing. doi:10.1098/rspb.2022.0477.

Batstone, R.T., Carscadden, K.A., Afkhami, M.E., and Frederickson, M.E. 2018. Using niche breadth theory to explain generalization in mutualisms. Ecology **99**(5): 1039–1050.

Batstone, R.T., Lindgren, H., Allsup, C.M., Goralka, L.A., Riley, A.B., Grillo, M.A., Marshall-Colon, A., and Heath, K.D. 2022b. Genome-Wide Association Studies across Environmental and Genetic Contexts Reveal Complex Genetic Architecture of Symbiotic Extended Phenotypes. mBio **13**(6). American Society for Microbiology. doi:10.1128/mbio.01823-22.

Bauer, W.D. 1981. Infection of Legumes by Rhizobia. Annu Rev Plant Physiol **32**(1): 407–449. doi:10.1146/annurev.pp.32.060181.002203.

Bolin, L.G., Lennon, J.T., and Lau, J.A. 2023. Traits of soil bacteria predict plant responses to soil moisture. Ecology **104**(2). doi:10.1002/ecy.3893.

Bosshard, P.P., Zbinden, R., and Altwegg, M. 2002. Paenibacillus turicensis sp. nov., a novel bacterium harbouring heterogeneities between 16S rRNA genes. Int J Syst Evol Microbiol **52**(6): 2241–2249. doi:10.1099/00207713-52-6-2241.

Brown, S., Grillo, M., Podowski, J., and Heath, K.D. 2020. Soil origin and plant genotype structure distinct microbiome compartments in the model legume Medicago truncatula. Microbiome **8**(139). doi:10.21203/rs.3.rs-21175/v1.

Burghardt, L.T., and diCenzo, G.C. 2023, April 1. The evolutionary ecology of rhizobia: multiple facets of competition before, during, and after symbiosis with legumes. Elsevier Ltd. doi:10.1016/j.mib.2023.102281.

Caetano-Anollés, G., and Gresshoff, P.M. 1991. Plant genetic control of nodulation. Annu Rev Microbiol **45**(1): 345–382. doi:10.1146/annurev.mi.45.100191.002021.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods **13**(7): 581–583. doi:10.1038/nmeth.3869.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. 2009. BLAST+: architecture and applications. BMC Bioinformatics **10**(1): 421. doi:10.1186/1471-2105-10-421.

Egamberdieva, D., Berg, G., Lindström, K., and Räsänen, L.A. 2010. Co-inoculation of Pseudomonas spp. with Rhizobium improves growth and symbiotic performance of fodder galega (Galega orientalis Lam.). Eur J Soil Biol **46**(3–4): 269–272. doi:10.1016/j.ejsobi.2010.01.005.

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S., and Pascal, G. 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. Bioinformatics **34**(8): 1287–1294. doi:10.1093/bioinformatics/btx791.

Fox, S.L., O'Hara, G.W., and Bräu, L. 2011. Enhanced nodulation and symbiotic effectiveness of Medicago truncatula when co-inoculated with Pseudomonas fluorescens WSM3457 and Ensifer (Sinorhizobium) medicae WSM419. Plant Soil **348**(1–2): 245–254. doi:10.1007/s11104-011-0959-8.

Garnier, S., Ross, N., Rudis, R., Camargo, A.P., Sciaini, M., and Scherer, C. 2023. viridis(Lite) - Colorblind-Friendly Color Maps for R.

Haney, C.H., Samuel, B.S., Bush, J., and Ausubel, F.M. 2015. Associations with rhizosphere bacteria can confer an adaptive advantage to plants. Nat Plants **1**(6): 15051. doi:10.1038/nplants.2015.51.

Hayden, H.S., Joshi, S., Radey, M.C., Vo, A.T., Forsberg, C., Morgan, S.J., Waalkes, A., Holmes, E.A., Klee, S.M., Emond, M.J., Singh, P.K., and Salipante, S.J. 2022. Genome Capture Sequencing Selectively Enriches Bacterial DNA and Enables Genome-Wide Measurement of Intrastrain Genetic Diversity in Human Infections. mBio **13**(5). doi:10.1128/mbio.01424-22.

Heath, K.D., and Tiffin, P. 2007. Context dependence in the coevolution of plant and rhizobial mutualists. Proceedings of the Royal Society B: Biological Sciences **274**(1620): 1905–1912. Royal Society. doi:10.1098/rspb.2007.0495.

Heath, K.D., and Tiffin, P. 2009. Stabilizing mechanisms in a legume-rhizobium mutualism. Evolution (N Y) **63**(3): 652–662. doi:10.1111/j.1558-5646.2008.00582.x.

Herridge, D.F., Peoples, M.B., and Boddey, R.M. 2008a. Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil **311**(1–2): 1–18. doi:10.1007/s11104-008-9668-3.

Herridge, D.F., Peoples, M.B., and Boddey, R.M. 2008b. Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil **311**(1–2): 1–18. doi:10.1007/s11104-008-9668-3.

Huse, S.M., Welch, D.M., Morrison, H.G., and Sogin, M.L. 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol **12**(7): 1889–1898. doi:10.1111/j.1462-2920.2010.02193.x.

Johnson, N.C., Graham, J., and Smith, F.A. 1997. Functioning of mycorrhizal associations along the mutualism–parasitism continuum*. New Phytologist **135**(4): 575–585. doi:10.1046/j.1469-8137.1997.00729.x.

Jones, E.I., Afkhami, M.E., Akçay, E., Bronstein, J.L., Bshary, R., Frederickson, M.E., Heath, K.D., Hoeksema, J.D., Ness, J.H., Pankey, M.S., Porter, S.S., Sachs, J.L., Scharnagl, K., and Friesen, M.L. 2015, November 1. Cheaters must prosper: Reconciling theoretical and empirical perspectives on cheating in mutualism. Blackwell Publishing Ltd. doi:10.1111/ele.12507.

Kent, A.D., and Triplett, E.W. 2002. Microbial Communities and Their Interactions in Soil and Rhizosphere Ecosystems. Annu Rev Microbiol **56**(1): 211–236. doi:10.1146/annurev.micro.56.012302.161120.

Klein, M., Stewart, J.D., Porter, S.S., Weedon, J.T., and Kiers, E.T. 2021. Evolution of manipulative microbial behaviors in the rhizosphere. John Wiley and Sons Inc. doi:10.1111/eva.13333.

Klein, M., Stewart, J.D., Porter, S.S., Weedon, J.T., and Kiers, E.T. 2022. Evolution of manipulative microbial behaviors in the rhizosphere. Evol Appl **15**(10): 1521–1536. doi:10.1111/eva.13333.

Kolde, R. 2019. pheatmap: Pretty Heatmaps.

Küster, H. 2013. Medicago truncatula. *In* Brenner's Encyclopedia of Genetics: Second Edition. Elsevier Inc. pp. 335–337. doi:10.1016/B978-0-12-374984-0.00915-3.

Kuzmanović, N., Fagorzi, C., Mengoni, A., Lassalle, F., and diCenzo, G.C. 2022. Taxonomy of Rhizobiaceae revisited: proposal of a new framework for genus delimitation. Int J Syst Evol Microbiol **72**(3). doi:10.1099/ijsem.0.005243.

de la Fuente Cantó, C., Simonin, M., King, E., Moulin, L., Bennett, M.J., Castrillo, G., and Laplaze, L. 2020. An extended root phenotype: the rhizosphere, its formation

and impacts on plant fitness. The Plant Journal **103**(3): 951–964. doi:10.1111/tpj.14781.

Lau, J.A., and Lennon, J.T. 2012. Rapid responses of soil microorganisms improve plant fitness in novel environments. Proceedings of the National Academy of Sciences **109**(35): 14058–14062. doi:10.1073/pnas.1202319109.

Love, M.I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**(12): 550. doi:10.1186/s13059-014-0550-8.

Martínez-Hidalgo, P., Galindo-Villardón, P., Igual, J.M., and Martínez-Molina, E. 2014. Micromonospora from nitrogen fixing nodules of alfalfa (Medicago sativa L.). A new promising Plant Probiotic Bacteria. Sci Rep **4**. Nature Publishing Group. doi:10.1038/srep06389.

Martínez-Hidalgo, P., and Hirsch, A.M. 2017. The nodule microbiome: N2fixing rhizobia do not live alone. American Phytopathological Society. doi:10.1094/PBIOMES-12-16-0019-RVW.

McMurdie, P.J., and Holmes, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One **8**(4): e61217. doi:10.1371/journal.pone.0061217.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., and Kondorosi, E. 2006. Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium* – legume symbiosis. Proceedings of the National Academy of Sciences **103**(13): 5230–5235. doi:10.1073/pnas.0600912103.

Nechayeva, A., Boyarshin, K., Bespalova, O., Iatsenko, V., Seliverstov, E., Klyueva, V., and Makanina, O. 2021. Intraspecies variability of the 16S rRNA gene of the soil bacteria *Acinetobacter lwoffii* and *Paenibacillus taichungensis*. BIO Web Conf **40**: 01009. doi:10.1051/bioconf/20214001009.

Neuwirth, E. 2022. RColorBrewer: ColorBrewer Palettes.

Nikodemova, M., Holzhausen, E.A., Deblois, C.L., Barnet, J.H., Peppard, P.E., Suen, G., and Malecki, K.M. 2023. The effect of low-abundance OTU filtering methods on the reliability and variability of microbial composition assessed by 16S rRNA amplicon sequencing. Front Cell Infect Microbiol **13**. doi:10.3389/fcimb.2023.1165295.

Nübel, U., Engelen, B., Felske, A., Snaidr, J., Wieshuber, A., Amann, R.I., Ludwig, W., and Backhaus, H. 1996. Sequence heterogeneities of genes encoding 16S rRNAs in Paenibacillus polymyxa detected by temperature gradient gel electrophoresis. J Bacteriol **178**(19): 5636–5643. doi:10.1128/jb.178.19.5636-5643.1996.

O'Brien, A.M., Jack, C.N., Friesen, M.L., and Frederickson, M.E. 2021. Whose trait is it anyways? Coevolution of joint phenotypes and genetic architecture in mutualisms. Proceedings of the Royal Society B: Biological Sciences **288**(1942): 20202483. doi:10.1098/rspb.2020.2483.

Paulson, J.N., Stine, O.C., Bravo, H.C., and Pop, M. 2013. Differential abundance analysis for microbial marker-gene surveys. Nat Methods **10**(12): 1200–1202. doi:10.1038/nmeth.2658.

Porter, S.S., Chang, P.L., Conow, C.A., Dunham, J.P., and Friesen, M.L. 2017. Association mapping reveals novel serpentine adaptation gene clusters in a

population of symbiotic Mesorhizobium. ISME J **11**(1): 248–262. doi:10.1038/ismej.2016.88.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res **41**(D1): D590–D596. doi:10.1093/nar/gks1219.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria. Available from https://www.R-project.org/.

Rahal, S., and Chekireb, D. 2021. Diversity of rhizobia and non-rhizobia endophytes isolated from root nodules of Trifolium sp. growing in lead and zinc mine site Guelma, Algeria. Arch Microbiol **203**(7): 3839–3849. doi:10.1007/s00203-021-02362-y.

Rahman, A., Manci, M., Nadon, C., Perez, I.A., Farsamin, W.F., Lampe, M.T., Le, T.H., Torres Martínez, L., Weisberg, A.J., Chang, J.H., and Sachs, J.L. 2023. Competitive interference among rhizobia reduces benefits to hosts. Current Biology **33**(14): 2988-3001.e4. doi:10.1016/j.cub.2023.06.081.

Rajendran, G., Patel, M.H., and Joshi, S.J. 2012. Isolation and characterization of nodule-associated Exiguobacterium sp. from the root nodules of fenugreek (Trigonella foenum-graecum) and their possible role in plant growth promotion. Int J Microbiol. doi:10.1155/2012/693982.

Riley, A.B., Grillo, M.A., Epstein, B., Tiffin, P., and Heath, K.D. 2023. Discordant population structure among rhizobium divided genomes and their legume hosts. Mol Ecol **32**(10): 2646–2659. John Wiley and Sons Inc. doi:10.1111/mec.16704.

Rosier, A., Beauregard, P.B., and Bais, H.P. 2021. Quorum Quenching Activity of the PGPR Bacillus subtilis UD1022 Alters Nodulation Efficiency of Sinorhizobium meliloti on Medicago truncatula. Front Microbiol **11**. Frontiers Media S.A. doi:10.3389/fmicb.2020.596299.

Sprent, J. 1987. The ecology of the nitrogen cycle. Cambridge University Press.

Tokgöz, S., Lakshman, D.K., Ghozlan, M.H., Pinar, H., Roberts, D.P., and Mitra, A. 2020. Soybean Nodule-Associated Non-Rhizobial Bacteria Inhibit Plant Pathogens and Induce Growth Promotion in Tomato. Plants **9**(11): 1494. doi:10.3390/plants9111494.

Tsiknia, M., Tsikou, D., Papadopoulou, K.K., and Ehaliotis, C. 2020. Multi-species relationships in legume roots: From pairwise legume-symbiont interactions to the plant – microbiome - soil continuum. FEMS Microbiol Ecol. doi:10.1093/femsec/fiaa222.

Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H., and Tilman, D.G. 1997. Human alteration of the global nitrogen cycle: sources and consequences. Ecological applications **7**(3): 737–750.

Wagner, M.R., Lundberg, D.S., Coleman-Derr, D., Tringe, S.G., Dangl, J.L., and Mitchell-Olds, T. 2014. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild Arabidopsis relative. Ecol Lett **17**(6): 717–726. doi:10.1111/ele.12276.

Wagner, M.R., Lundberg, D.S., del Rio, T.G., Tringe, S.G., Dangl, J.L., and Mitchell-Olds, T. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. Nat Commun **7**(1): 12151. doi:10.1038/ncomms12151.

Wendlandt, C.E., Regus, J.U., Gano-Cohen, K.A., Hollowell, A.C., Quides, K.W., Lyu, J.Y., Adinata, E.S., and Sachs, J.L. 2019. Host investment into symbiosis varies among genotypes of the legume Acmispon strigosus, but host sanctions are uniform. New Phytologist **221**(1): 446–458. Blackwell Publishing Ltd. doi:10.1111/nph.15378.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer International Publishing, New York, NY.

Wright, E.S. 2016. Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. R J **8**(1): 352. doi:10.32614/RJ-2016-025.

Zilles, J.L., Rodríguez, L.F., Bartolerio, N.A., and Kent, A.D. 2016. Microbial community modeling using reliability theory. ISME J **10**(8): 1809–1814. doi:10.1038/ismej.2016.1.

FIGURES

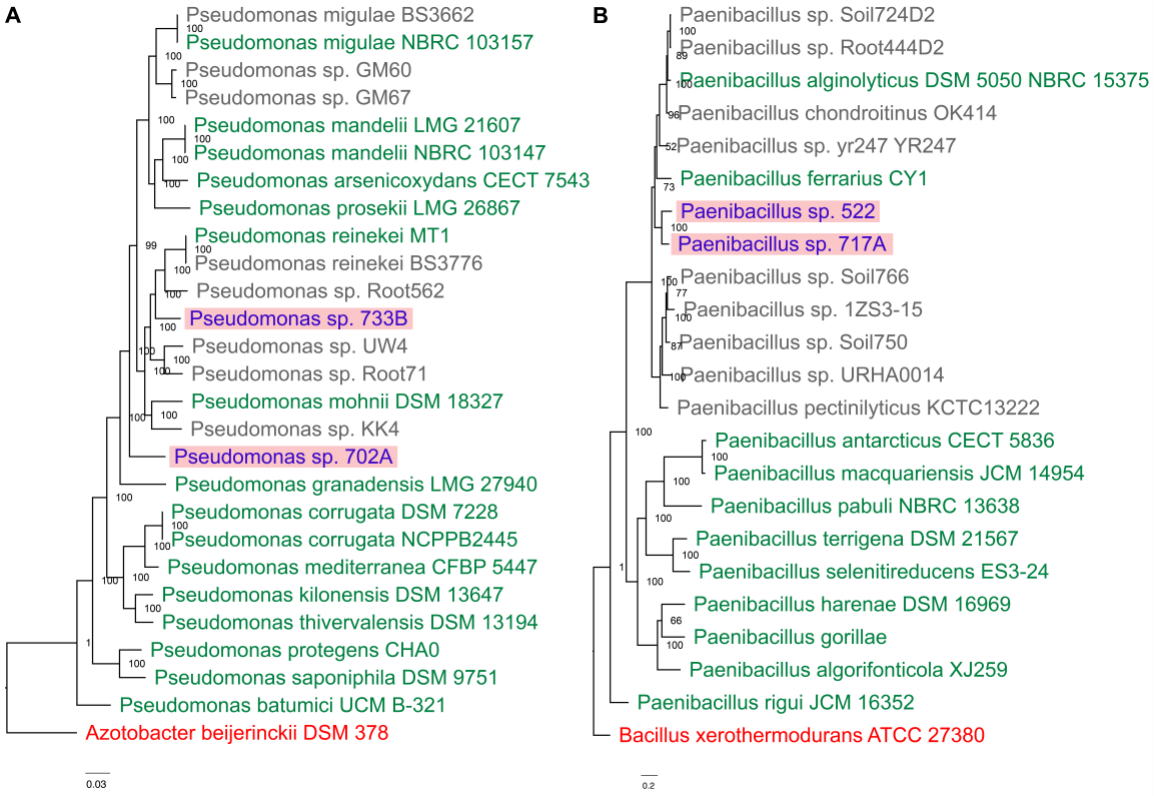


Figure 1. Novel *Paenibacillus* and *Pseudomonas* species among non-rhizobia endophytes. Multi-locus *Pseudomonas* (A) and *Paenibacillus* (B) species trees constructed with concatenated genome alignment matrices by AutoMLST (Alanjary et al., 2019). Blue taxa highlighted in red are query strains, green taxa are type strains, and grey taxa are non-type strains. The outgroups are colored in red. Phylogenies were inferred using IQ-TREE (Nguyen et al., 2015) with ModelFinder (Kalyaanamoorthy et al., 2017). Bootstrap support percentages for 1000 ultrafast bootstraps (Minh et al., 2013) are shown at each node. A scale for genome distance (Ondov et al., 2016) is provided for each tree.

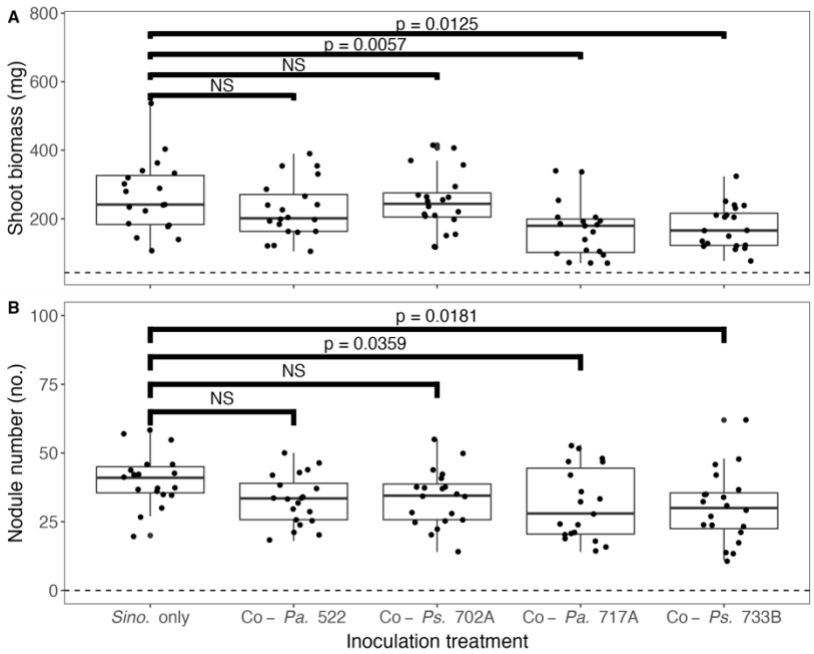


Figure 2. Co-inoculations of *Sinorhizobium meliloti* and *Paenibacillus* sp. 717A or *Pseudomonas sp.* 733B indirectly reduced host shoot mass and the number of nodules formed. Co-inoculations with *S. meliloti* and *Paenibacillus* sp. 522 or *Pseudomonas* sp. 702A did not yield different shoot masses (A) or numbers of nodules (B) than plants inoculated with *S. meliloti* only ($P > 0.05$, type II ANOVA, FDR adjusted), while co-inoculations with *Paenibacillus* sp. 717A or *Pseudomonas* sp. 733B significantly reduced the average shoot mass (A) and number of nodules (B) on plants ($P < 0.01$, $P < 0.05$, type II ANOVA, FDR adjusted). *Sino.* only: *S. meliloti* 141 single-inoculation; Co – *Pa.* 522: *S. meliloti* 141 and *Paenibacillus* sp. 522 co-inoculation; Co – *Ps.* 702A: *S. meliloti* 141 and *Pseudomonas* sp. 702A co-inoculation; Co – *Pa.* 717A: *S. meliloti* 141 and *Paenibacillus* sp. 717A co-inoculation; Co – *Ps.* 733B: *S. meliloti* 141 and *Pseudomonas* sp. 733B co-inoculation; NS: not significant.

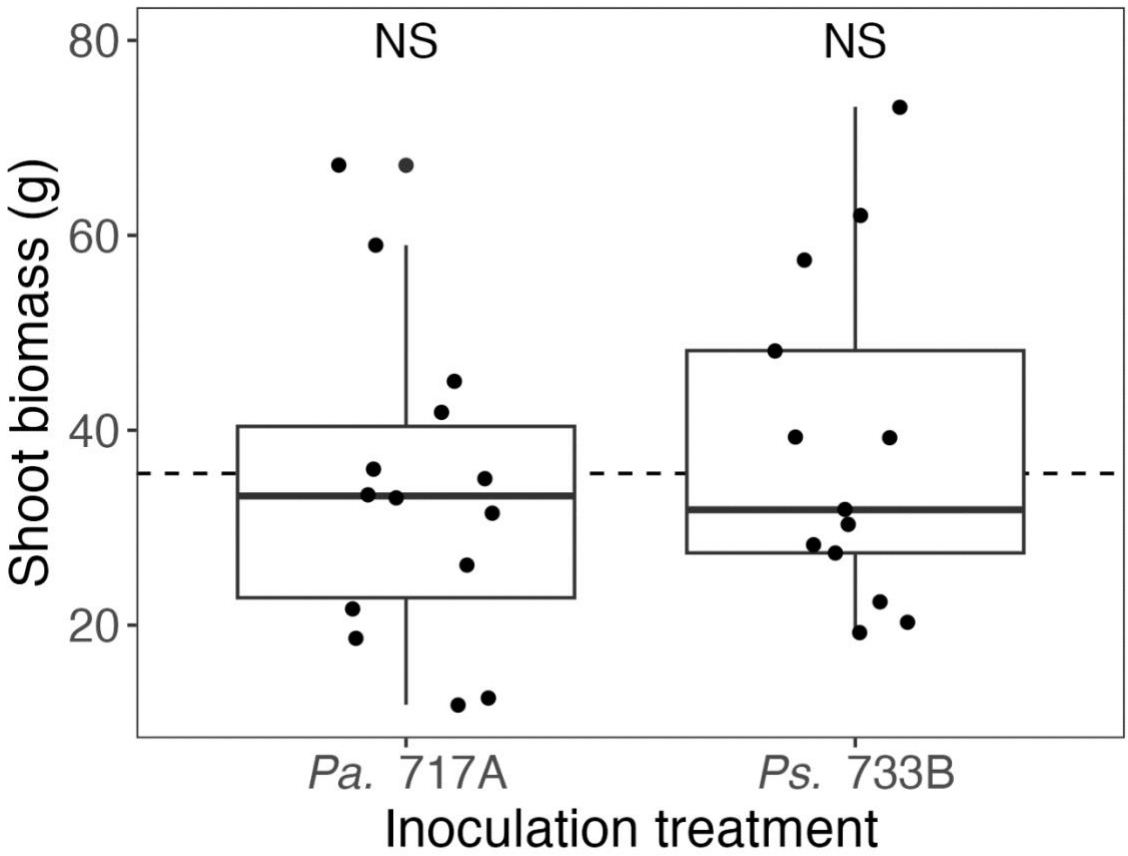


Figure 3. NREs *Paenibacillus* sp. 717A and *Pseudomonas* sp. 733B had no direct effects on nitrogen-supplemented plants. Average shoot biomass for nitrogen-supplemented (NH_4NO_3) plants in the direct effects of NREs experiment are given. Inoculation treatment did not have a significant effect on shoot mass ($P > 0.1$, type II ANOVA). *Pa.* 717A: *Paenibacillus* sp. 717A single inoculation; *Ps.* 733B *Pseudomonas* sp. 733B single inoculation; NS: not significant.

Figure 4. Amplicon sequence variants inferred from nodule and root samples of *Medicago truncatula*. Amplicon sequence variant (ASV) counts transformed by variance-stabilization are given in the heatmap matrix. Columns correspond to individual DNA libraries sequenced and extraced from *M. truncatula* samples across a combination

of inoculation treatments, tissue sections, and surface-sterilization treatments. Rows represent the abundance of individual ASVs in a sample. For simplicity, only the 50 most abundant ASVs are shown. Family and genus taxonomy labels are as obtained by DECIPHER (Wright 2016) using the SILVA SSU r138 (Quast et al. 2012) training set, NA: unassigned. AVS highlighted in red are ASVs of interest suspected to represent our inoculum strains.

SUPPLEMENTARY FIGURES

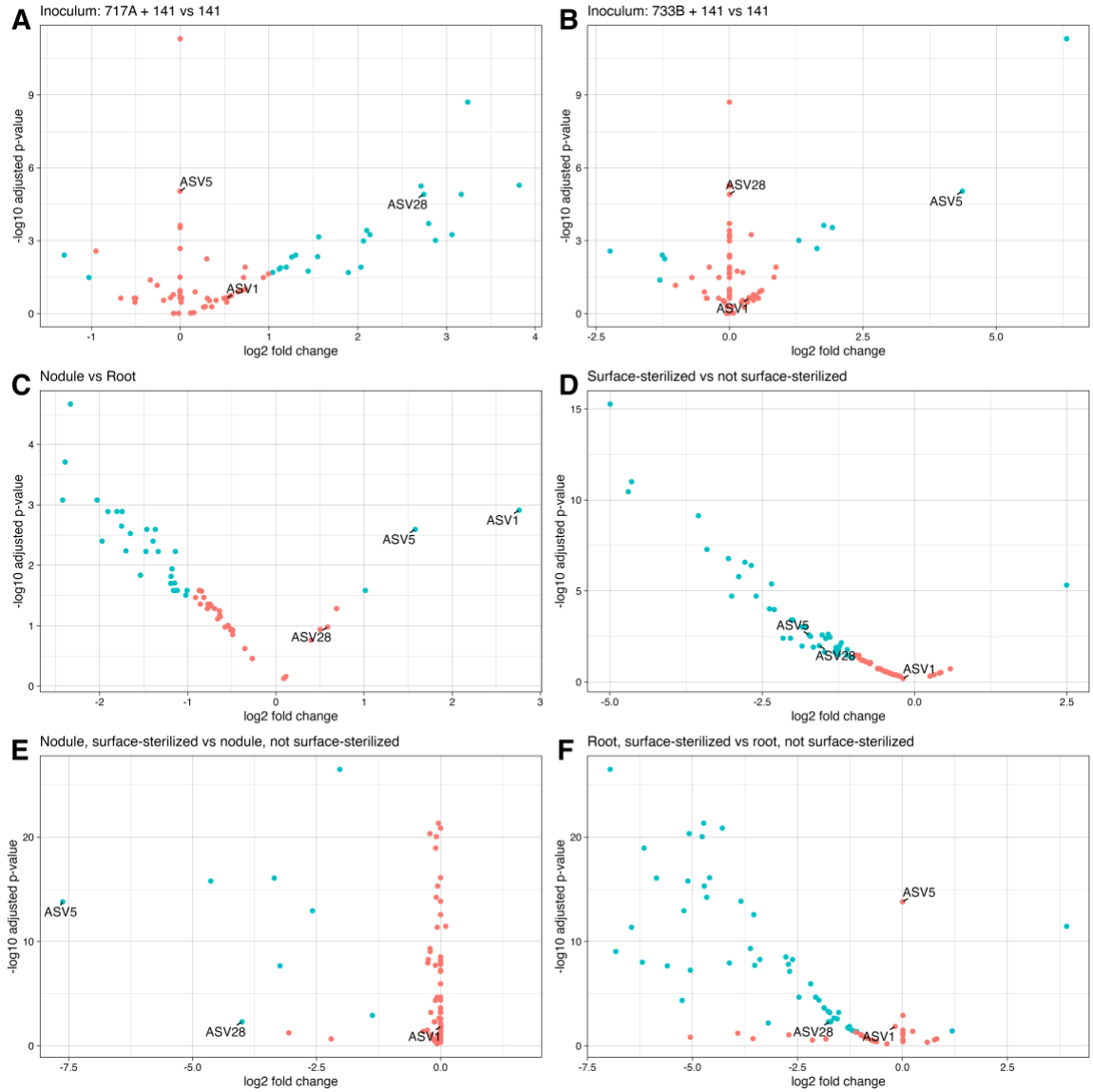


Figure S1. Differential abundance of ASVs across inocula, tissue sections, and surface-sterilziation treatments. Using a minimum fold change 1.5 and *P* (false-discovery rate adjusted) < 0.05, amplicon sequence variants (ASVs) diferentially abundant across contrasts of inocula, tissues, and surface-sterilization treatments were inferred. Points refer to individual ASVs, blue points correspond to signifiicant ASVs while red points are insignificant. The second term in plot titles indicate baselines in

comparisons. ASVs representing inocula strains are labeled, ASV1 = *S. meliloti* 141; ASV28 = *Paenibacillus* sp. 717A; ASV5 = *Pseudomonas* sp. 733B. (A) ASV28 was enriched in *Paenibacillus* sp. 717A + *S. meliloti* 141 co-inoculated samples compared to samples inoculated with *S. meliloti* 141 only. Similarly, (B) ASV5 was enriched in *Pseudomonas* sp. 733B + *S. meliloti* 141 co-inoculated samples compared to samples inoculated with *S. meliloti* 141 only. (C) Both ASV5 and ASV1 were enriched in nodule samples compared to root samples. (D) ASV5 and ASV28 were enriched in non-surface-sterilized samples compared to surface-sterilized samples, while ASV1 was not differentially abundant across either group. (E) Among just nodule samples, ASV5 and ASV28 were enriched in non-surface-sterilized samples while ASV1 was not differentially abundant. (F) Among just root samples, only ASV28 was enriched in non-surface-sterilzied root samples while ASV1 and ASV5 were not different across either group.

SUPPLEMENTARY TABLE LEGENDS

Table S1. List of genes accessed and concatenated by AutoMLST (Alanjary et al., 2019) for multi-locus sequence analysis of *Paenibacillus* strains 522 and 717A.

Table S2. List of genes accessed and concatenated by AutoMLST (Alanjary et al., 2019) for multi-locus sequence analysis of *Pseudomonas* strains 702A and 733B.

Table S3. Original complete set of mash distances (Ondov et al., 2016) generated by AutoMLST (Alanjary et al., 2019) for multi-locus sequence analysis of *Paenibacillus* strains 522 and 717A.

Table S4. Set of mash distances (Ondov et al., 2016) generated by AutoMLST (Alanjary et al., 2019) for multi-locus sequence analysis of *Pseudomonas* strains 702A and 733B.

Table S5. Metrics and corresponding values from OrthoANI (Lee et al., 2016) estimations comparing genomes of *Paenibacillus* species 522 and 717A as well as *Pseudomonas* species 702A and. 733B.

Table S6. Raw phenotypic data collected from the indirect effects greenhouse experiment.

Table S7. Estimated marginal means (Lenth, 2022) and false discovery rate adjusted p-values (Lenth, 2022) for every pairwise comparison between all treatments for each trait measured in the indirect effects greenhouse experiment.

Table S8. Raw phenotypic data collected from the direct effects nitrogen-addition experiment.

Table S9. Amplicon sequence variant (ASV) table generated by DADA2 (Callahan et al. 2016) from amplified 16S V3-V4 rRNA regions of DNA extracted from nodule and root tissues of co-inoculated and single-inoculated *M. truncatula* plants. Column names refer

to ASV names, row names are individual DNA samples. For sample metadata, see Table S13.

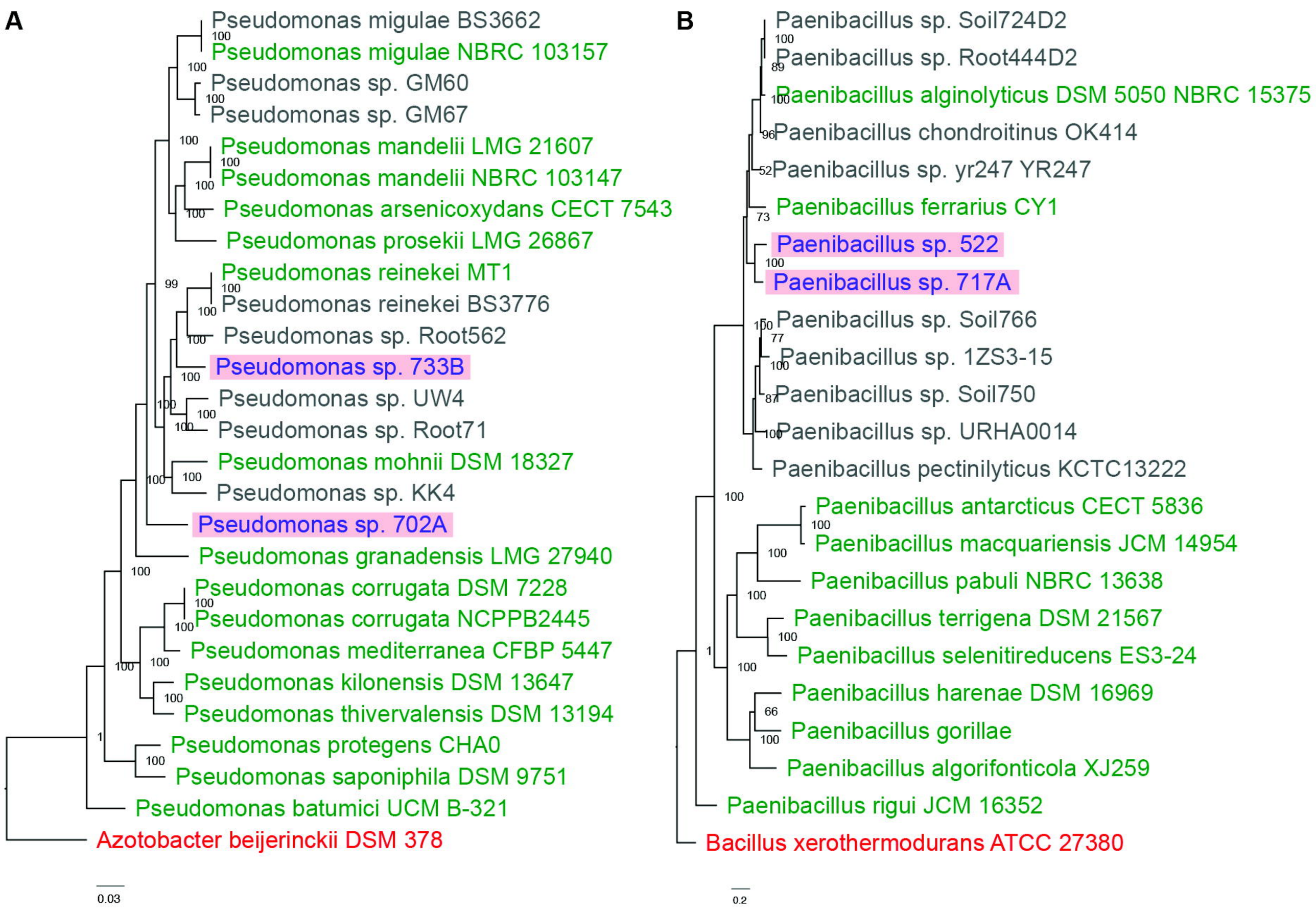
Table S10. Nucleotide sequences of ASVs present in Table S9.

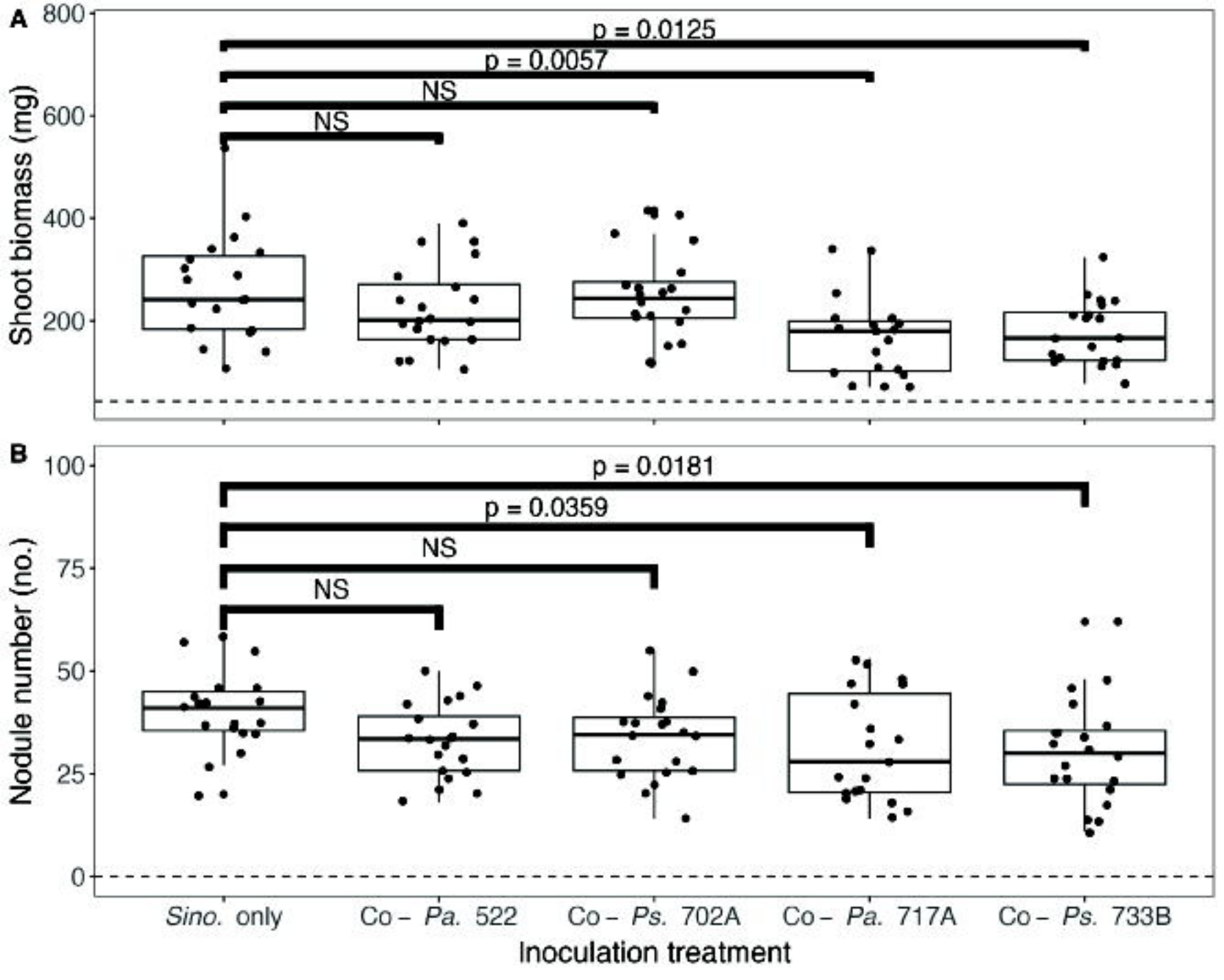
Table S11. Nucleotide BLAST (Camacho et al. 2009) of ASV sequences in Table S9 against 16S rRNA gene annotations from *Sinorhizobium meliloti* 141, *Paenibacillus* sp. 717A, and *Pseudomonas* sp. 733B. Minimum percent identity of 80% and maximum e-value of 0.01.

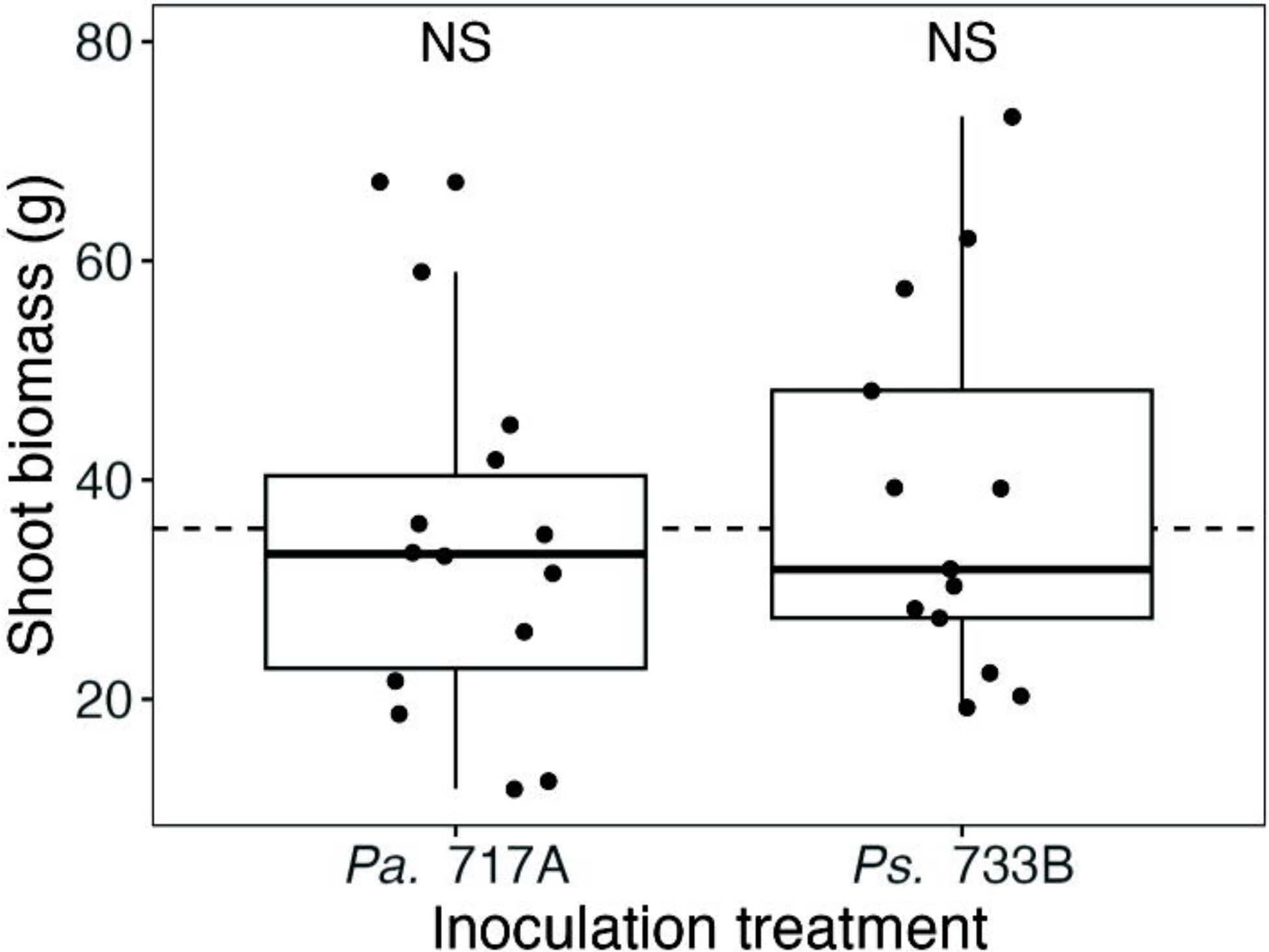
Table S12. Nucleotide BLAST (Camacho et al. 2009) of ASV sequences in Table S9 against the top 50 OTU sequences from Brown et al. (Brown et al. 2020). Minimum percent identity of 80% and maximum e-value of 0.01.

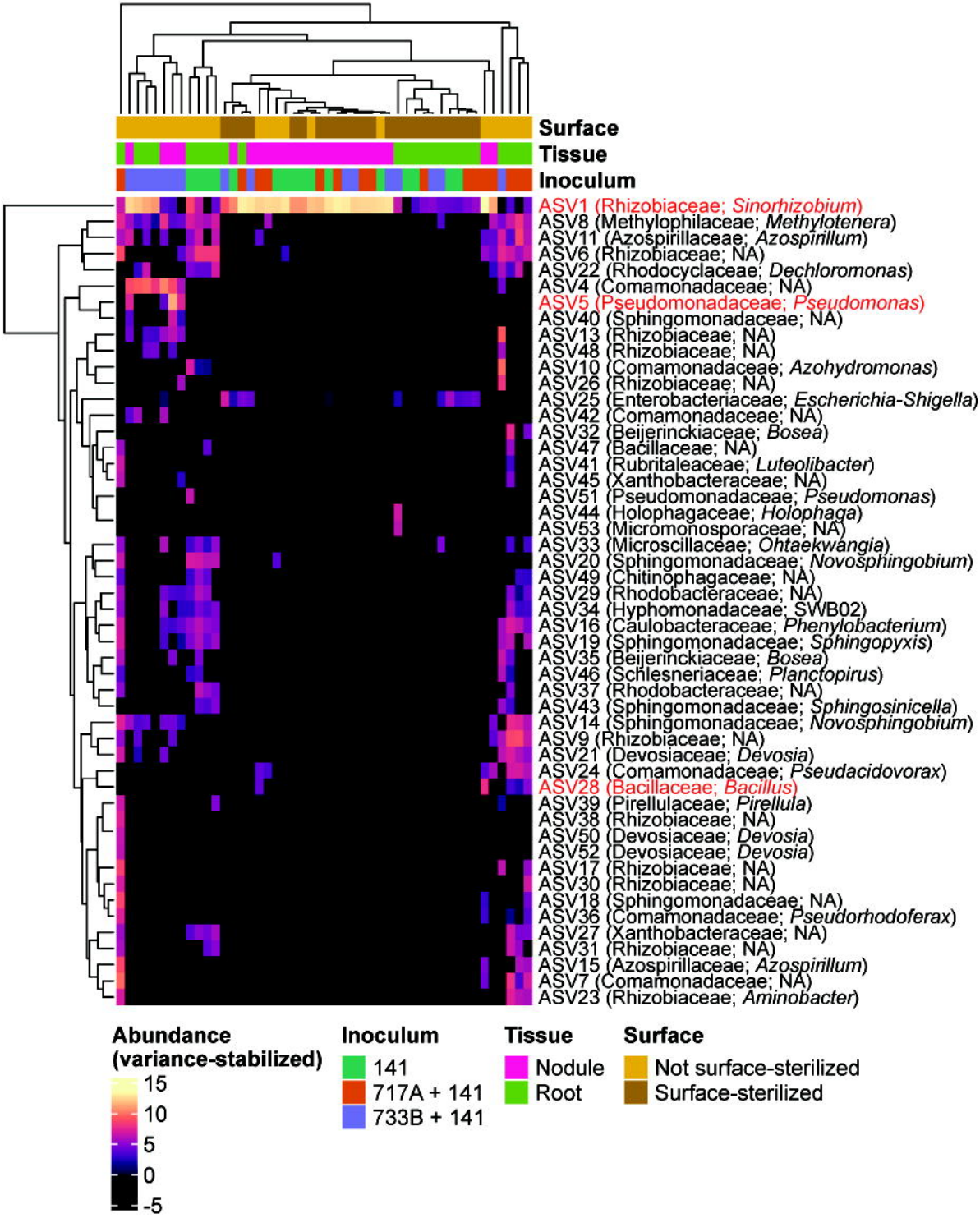
Table S13. Metadata for tissue occupancy experiment DNA samples as reported in Table S9.

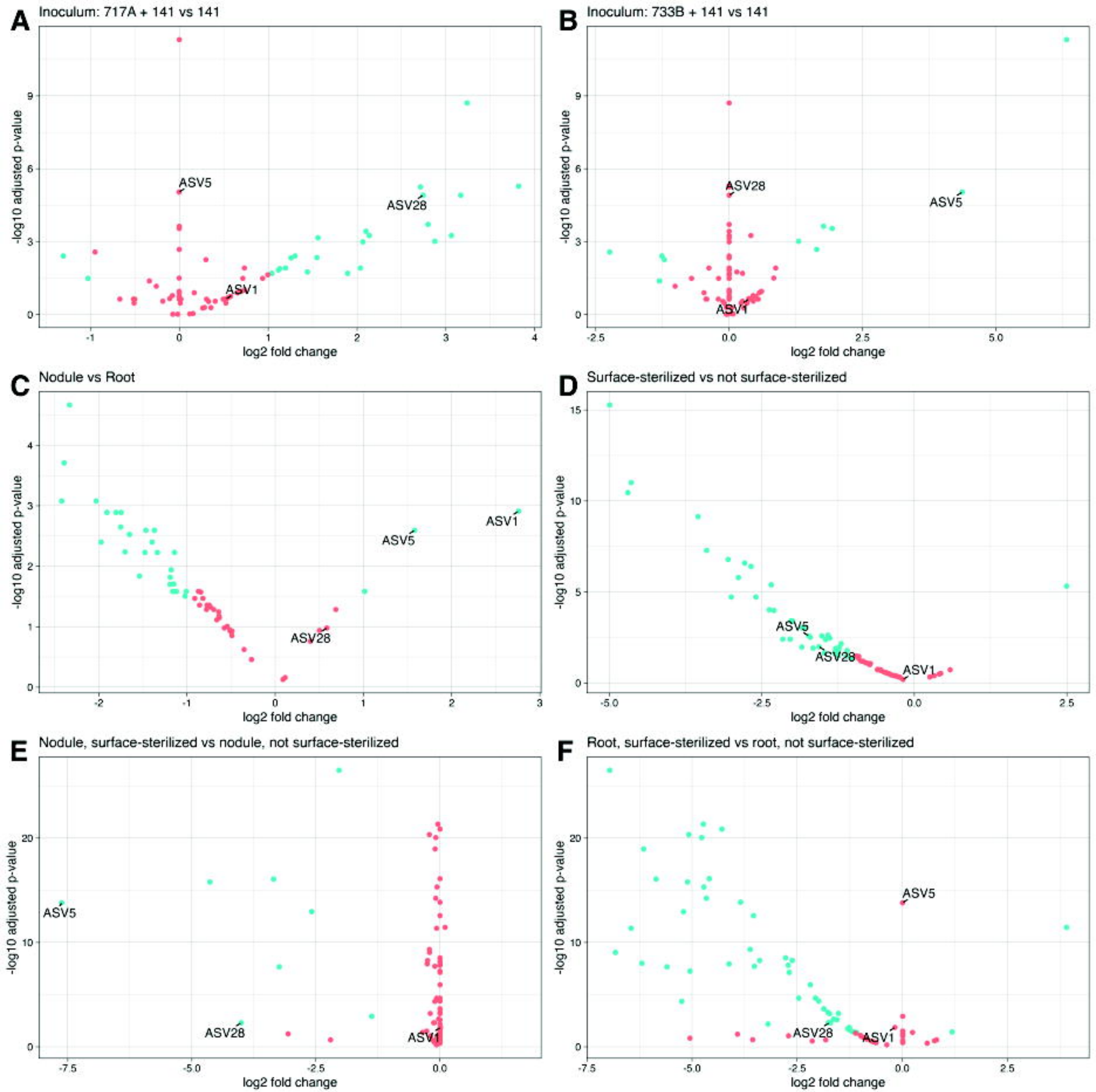
Table S14. Wald test results from DESeq2 (Love et al. 2014) comparing the difference in abundance for ASVs present in Table S9 across inoculum treatments, tissue sections, and surface-sterilization treatments.


A
Pseudomonas migulae BS3662
Pseudomonas migulae NBRC 103157
Pseudomonas sp. GM60
Pseudomonas sp. GM67
Pseudomonas mandelii LMG 21607
Pseudomonas mandelii NBRC 103147
Pseudomonas arsenicoxydans CECT 7543
Pseudomonas prosekii LMG 26867
Pseudomonas reinekei MT1
Pseudomonas reinekei BS3776
Pseudomonas sp. Root562
Pseudomonas sp. 733B
Pseudomonas sp. UW4
Pseudomonas sp. Root71
Pseudomonas mohnii DSM 18327
Pseudomonas sp. KK4
Pseudomonas sp. 702A
Pseudomonas granadensis LMG 27940
Pseudomonas corrugata DSM 7228
Pseudomonas corrugata NCPPB2445
Pseudomonas mediterranea CFBP 5447
Pseudomonas kilonensis DSM 13647
Pseudomonas thivervalensis DSM 13194
Pseudomonas protegens CHA0
Pseudomonas saponiphila DSM 9751
Pseudomonas batumici UCM B-321
Azotobacter beijerinckii DSM 378
0.03
B
Paenibacillus sp. Soil724D2
Paenibacillus sp. Root444D2
Paenibacillus alginolyticus DSM 5050 NBRC 15375
Paenibacillus chondroitinus OK414
Paenibacillus sp. yr247 YR247
Paenibacillus ferrarius CY1
Paenibacillus sp. 522
Paenibacillus sp. 717A
Paenibacillus sp. Soil766
Paenibacillus sp. 1ZS3-15
Paenibacillus sp. Soil750
Paenibacillus sp. URHA0014
Paenibacillus pectinilyticus KCTC13222
Paenibacillus antarcticus CECT 5836
Paenibacillus macquariensis JCM 14954
Paenibacillus pabuli NBRC 13638
Paenibacillus terrigena DSM 21567
Paenibacillus selenitireducens ES3-24
Paenibacillus harenae DSM 16969
Paenibacillus gorillae
Paenibacillus algorifonticola XJ259
Paenibacillus rigui JCM 16352
Bacillus xerothermodurans ATCC 27380
0.2


A
800
600
400
200
Shoot biomass (mg)
p = 0.0125
p = 0.0057
NS
NS
B
100
75
50
25
0
Nodule number (no.)
p = 0.0181
p = 0.0359
NS
NS
Sino. only
Co – Pa. 522
Co – Ps. 702A
Co – Pa. 717A
Co – Ps. 733B
Inoculation treatment


NS
NS
80
60
40
20
Shoot biomass (g)
Pa. 717A
Ps. 733B
Inoculation treatment


Surface
Tissue
Inoculum
ASV1 (Rhizobiaceae; Sinorhizobium)
ASV8 (Methylophilaceae; Methylotenera)
ASV11 (Azospirillaceae; Azospirillum)
ASV6 (Rhizobiaceae; NA)
ASV22 (Rhodocyclaceae; Dechloromonas)
ASV4 (Comamonadaceae; NA)
ASV5 (Pseudomonadaceae; Pseudomonas)
ASV40 (Sphingomonadaceae; NA)
ASV13 (Rhizobiaceae; NA)
ASV48 (Rhizobiaceae; NA)
ASV10 (Comamonadaceae; Azohydromonas)
ASV26 (Rhizobiaceae; NA)
ASV25 (Enterobacteriaceae; Escherichia-Shigella)
ASV42 (Comamonadaceae; NA)
ASV32 (Beijerinckiaceae; Bosea)
ASV47 (Bacillaceae; NA)
ASV41 (Rubritaleaceae; Luteolibacter)
ASV45 (Xanthobacteraceae; NA)
ASV51 (Pseudomonadaceae; Pseudomonas)
ASV44 (Holophagaceae; Holophaga)
ASV53 (Micromonosporaceae; NA)
ASV33 (Microscillaceae; Ohtaekwangia)
ASV20 (Sphingomonadaceae; Novosphingobium)
ASV49 (Chitinophagaceae; NA)
ASV29 (Rhodobacteraceae; NA)
ASV34 (Hyphomonadaceae; SWB02)
ASV16 (Caulobacteraceae; Phenylobacterium)
ASV19 (Sphingomonadaceae; Sphingopyxis)
ASV35 (Beijerinckiaceae; Bosea)
ASV46 (Schlesneriaceae; Planctopirus)
ASV37 (Rhodobacteraceae; NA)
ASV43 (Sphingomonadaceae; Sphingosinicella)
ASV14 (Sphingomonadaceae; Novosphingobium)
ASV9 (Rhizobiaceae; NA)
ASV21 (Devosiaceae; Devosia)
ASV24 (Comamonadaceae; Pseudacidovorax)
ASV28 (Bacillaceae; Bacillus)
ASV39 (Pirellulaceae; Pirellula)
ASV38 (Rhizobiaceae; NA)
ASV50 (Devosiaceae; Devosia)
ASV52 (Devosiaceae; Devosia)
ASV17 (Rhizobiaceae; NA)
ASV30 (Rhizobiaceae; NA)
ASV18 (Sphingomonadaceae; NA)
ASV36 (Comamonadaceae; Pseudorhodoferax)
ASV27 (Xanthobacteraceae; NA)
ASV31 (Rhizobiaceae; NA)
ASV15 (Azospirillaceae; Azospirillum)
ASV7 (Comamonadaceae; NA)
ASV23 (Rhizobiaceae; Aminobacter)
Abundance (variance-stabilized)
15
10
5
0
-5
Inoculum
141
717A + 141
733B + 141
Tissue
Nodule
Root
Surface
Not surface-sterilized
Surface-sterilized


A
Inoculum: 717A + 141 vs 141
-log10 adjusted p-value
log2 fold change
ASV5
ASV28
ASV1
B
Inoculum: 733B + 141 vs 141
-log10 adjusted p-value
log2 fold change
ASV28
ASV5
ASV1
C
Nodule vs Root
-log10 adjusted p-value
log2 fold change
ASV1
ASV5
ASV28
D
Surface-sterilized vs not surface-sterilized
-log10 adjusted p-value
log2 fold change
ASV5
ASV28
ASV1
E
Nodule, surface-sterilized vs nodule, not surface-sterilized
-log10 adjusted p-value
log2 fold change
ASV5
ASV28
ASV1
F
Root, surface-sterilized vs root, not surface-sterilized
-log10 adjusted p-value
log2 fold change
ASV5
ASV28
ASV1